Incorporating Different Frameworks to Interpret Imperfect Recall

Ruochen Dong

Submitted in fulfilment of the requirements of the Degree of PhD

Adam Smith Business School, College of Social Science
University of Glasgow

August 2022

Abstract

When computing the expected payoffs of a player at an information set with absentmindedness, there are many issues regarding the belief system that the player forms on the probability measures over histories in that information set. One of those is the issue raised by the circle “strategy- belief- strategy". The circle “strategy- belief- strategy" indicates the belief at an information set is formed by the strategy, and the decision maker needs to choose an optimal strategy in the premise of the belief, which implies a problematic logic. To solve this problem, we interpret a finite decision problem with absentmindedness into a psychological multiself game form. A self called planner represents the decision maker before the decision problem starts. A doer represents the decision maker at each node during execution. The planner could choose a full strategy for the decision problem based on his different beliefs on how the doers behave. A confident planner believes the doers will follow his chosen strategy. A knowledgable planner believes the doers will reevaluate the decision problem themselves and choose conditional optimal one-shot behaviour. The doers behave the same way as described in the knowledgable planner’s belief. Thus, we transfer a decision problem with absentmindedness into a two-stage psychological game. The existence of a psychological multiself equilibrium with a confident planner indicates that an ex-ante optimal strategy is not only modified multiself consistent (defined in Piccione & Rubinstein (1997a)) but stable in terms of multiple one-shot deviations. It indicates if doers do not cooperate, they cannot reach a better situation than the situation if they execute the ex-ante optimal strategy. The existence of a psychological multiself equilibrium with a knowledgable planner presents the situation if the planner expects potential deviations.

In conventional decision theory, it is not much different whether a decision maker knows future imperfect recall, except there is an extra requirement of strategy choose if the decision maker knows. It is interesting to explore how a decision maker behaves if he knows he has imperfect recall and prepares for it in the way that he takes how he will evaluate the decision problem at the future information set into consideration when he reassesses the decision problem currently. We call such a decision maker sophisticated. We develop the sophisticated recursive calculation rule to describe the process of reconsideration at an information set. The resulted expected payoff function is called IS expected payoff function. The IS expected payoff function is used to examine the time consistency of a strategy at an information set in a decision problem and the sequential rationality of a strategy profile at a collection of information sets in a game. Then, it concludes a strategy is ex-ante optimal if and only if it is IS time consistent and an IS sequential equilibrium exists in any finite game.

Contents

Abstract **1**

Introduction **7**

**1 Psychological Multiself Approach:
How Does a Decision Maker with Absentmindedness React to Time Inconsistency?** **20**

1.1 Introduction 20

1.2 Previous Works of Decision Problems with Absentmindedness 29

1.3 Manager-agents Model 32

1.4 All about Psychological Games 36

1.4.1 Psychological Game Framework 37

1.4.2 Fairness and Reciprocity 39

1.4.3 Guilt Aversion 41

1.4.4 Brief Literature on Other Issues 42

1.4.5 Why We Incorporate Psychological Games into Our Approach? 44

1.5 Revisit Absentminded Driver Paradox 46

1.5.1 Standard Absentminded Driver Paradox 47

1.5.2 $x-y-z$ General Version of Absentminded Driver Paradox 50

1.5.3 Recap Equilibrium Results 53

1.6 Psychological Multiself Approach 56

1.6.1 Basic Definitions and Notation 56

1.6.2 Psychological Multiself Equilibrium 64

1.6.3 Multiple Multiselves Consistency 79

1.6.4 "Time Consistency" Issue in Psychological Multiself Approach with a Knowledgable Planner 85

1.7 Conclusions and Extension 87

2 Analysis of Decision Problems with Imperfect Recall Using IS Expected Payoff Function 90

2.1 Introduction 90

2.2 Imperfect Recall and Time Inconsistency 94

2.2.1 Time Inconsistency in Imperfect Recall (Examples) 95

2.2.2 Time Inconsistency in Imperfect Recall (General Explanation) 97

2.2.3 Difficulties in Interpretation with the Presence of Imperfect Recall 99

2.3 Literature Review 107

2.4 Necessary Features of the Definition 110

2.5 Expected Payoffs at an Information Set 114

2.5.1 Two Examples 118

2.5.2 Calculation Rules Applied to Decision Problems with Perfect Recall 120

2.5.3 Sophisticated Recursive Calculation Rule 125

2.5.4 Alternative Formation of the Sophisticated Recursive Calculation Rule 133

2.5.5 Cross-branch Information Sets 134

2.6 IS-time Consistency and One Information Set Deviation Property 136

2.7 Strongly IS-time Consistency 145

2.8 Reconsideration at Multiple Information Sets 146

2.9 Conclusions and Discussion 151

2.10 Appendix 153

3 Sequential Equilibrium in Games with Imperfect Recall 162

3.1 Introduction 162

3.2 Relevant Literatures 166

3.3 What Is a Player? 168

3.4 What Is a Strategy? 170

3.5 Notations and Definitions 173

3.6 IS Expected Payoff Function 177

3.6.1 IS Expected Payoff Function at an Information Set 178

3.6.2 Discussion on Deviations at Multiple Information Sets 181

3.7 IS Sequential Equilibrium and Different Solution Concepts 184

3.7.1 An Example 184

3.7.2 Switch of a Strategy during Execution 186

3.7.3 Perturbed Games 188

3.7.4 Perfect Equilibrium 189

3.7.5 IS Sequential Equilibrium 191

3.7.6 IS Nash Equilibrium, Nash Equilibrium and Agent Equilibrium 194

3.8 Conclusions and Discussions 198

3.9 Appendix 199

Conclusions and Discussion **204**

Appendix **207**

Acknowledgement **217**

Declaration **218**

Introduction

In classic game theory, players are assumed to be perfectly rational. Requirements for perfect rationality are that players are perfect in memory, always choose the optimal strategy according to their payoff function, never make mistakes, and so on. However, in the real world, individuals present imperfection and “irrationality" in most situations. For those “irrational" behaviours, we could say that probably the principle behind them has not been found which makes people’s behaviours seem to be “irrational". For example, an individual might choose a decision driven by emotions such that his choice is not the optimal one from the viewpoint of the third party. We could apply psychological games to analyse this situation which makes the individual’s choice rationalised. As for imperfection, few previous works discuss imperfect recall. A player is assumed to have perfect recall, if he can remember all his past actions and all previously acquired information about the game and other players. A game (decision problem) is said to have perfect recall if all the players (the decision maker) are assumed to have perfect recall.

The reason why seldom researches on imperfect recall is that a game becomes rapidly more complicated to interpret once relax the assumption of perfect recall, especially absentmindedness presents. There are many issues which are discussed in conventional game theory about perfect recall but cannot be defined in a reasonable way in the situation of imperfect recall, such as the belief of probability distribution of being at each node in an information set with absentmindedness. There are many issues which are not important in the situation of perfect recall but become crucial when imperfect recall presents. For example, in a finite decision problem with perfect recall, it does not matter whether the decision maker chooses a strategy ex-ante or when he moves to each node. An ex-ante optimal strategy is time consistent. It illustrates that if the choice of a strategy seems to be rational before the decision problem starts, then during execution, the decision maker dose not have a better option and will not deviate from the ex-ante chosen strategy. However, when imperfect recall presents an ex-ante optimal strategy is not necessarily to be time consistent, such as the absentminded driver paradox. A decision maker might deviate from the ex-ante optimal strategy when he moves to some node even if the decision maker's preference does not change and no new information is known by the decision maker. In this case, when a decision maker makes decisions becomes significant. The result of a decision problem could be different if the decision maker chooses a strategy ex-ante or he chooses a strategy at every move during execution.

Some people might ask whether it is necessary to put so much effort to develop a reasonable framework for analysing imperfect recall since the games with imperfect recall seem to be much more complicated. We would like to answer yes to this question. There are many situations which should be described by imperfect recall rather than a traditional game.

Examples of Imperfect Recall

Ambrus-Lakatos (1999) provides several economic examples to show imperfect recall. The first one emphases the fact that the decision maker is assumed to choose the conditionally optimal one-shot behavioural rule, and it provides the motivation to model multself approaches in games (decision problems) with imperfect recall. The example is about election and government. When governments face elections, they would like to convince voters that the quality of their lives could be improved during their tenure. However, in most cases, things are not going on very well. Then, if the government finishes their first tenure and expects to be elected again in the next election. They might take the action of blaming other agents that they cause the failure. These other agents might be some institutions such that they almost work independently but are limited controlled by the government. Thus, the government could adopt the excuse that it is impossible to improve the country to a certain level due to the independence of those institutions. In fact, the real relationship between the government and other institutions is much ambiguous. The government could switch among different interpretations of the relationship before and after the election rapidly to promote themselves.

Another example Ambrus-Lakatos (1999) exposes is the bargaining between an employer and an employee about the amount of salary which should be paid to the employee. If they cannot agree to the same amount about the wage privately, they could choose to submit their claims to a court of arbitration. The arbitration has two stages. A first arbitrator studies the dispute and offers terms for a settlement. The employer and the employee could choose to accept the settlement or not. If not, they go to the second arbitrator and they must accept whatever settlement the second arbitrator provides. Assume The second arbitrator can only have access to the settlement proposed by the first arbitrator but not the information about why the first arbitrator offers such a settlement. Both arbitrators are motivated to make an impartial decision due to reasons such as reputation. If we regard the two stages with different arbitrators as the two sections of one agent, the agent losses the information regarding the game when it is executed in the second stage.

In recent several years, credit cards and online shopping have expanded rapidly. Poddar et al. (2015) proposes the use of credit cards might cause an individual is willing to pay more using credit card than cash for the same product. It is contrary to the decision theory in the conventional agent model. Poddar et al. (2015) present evidences that people are overspending when they use credit cards. One of those pieces of evidence is that the world bank collects data about the increase in income and the increase in credit card debt. The world bank found out that, in the past thirty years, the credit card debt increased near fifteen times while the per capita income increased about four times. The imperfect recall in this example is explained to be self-serving, and the revised memory leads them to believe that they were more successful and made better decisions in the past. Thus, unless people are reminded of what were the consequences of their prior choice of actions through a certain external device, the motives that led to these actions may not be remembered accurately and completely.

However, we would rather explain imperfect recall causes overspending by using credit cards as the way that people cannot remember all their past consumption since we cannot observe the reduction of saving visually and intuitively. Especially, if one has multiple credit cards at different banks, it becomes more difficult for him to track his consumption. Then, it might cause overspending. For example, assume that an individual makes the consumption plan at the start of every month. However, in the midmonth, the individual would like to choose a new consumption plan due to some external stimuli. At that time, some restrictions make the person unable to track his past consumption. For example, he might use the credit card multiple times in one day. Or there is a delay in the payment. Thus, when the individual makes the new plan in midmonth, he could be regarded to have imperfect recall about his past actions about consumption. If he would like to see whether there is another consumption plan which is better than the one he made at the beginning of this month, he is actually examining the time consistency of an ex-ante strategy. The fact of imperfect recall might motivate him to spend more using a credit card.

If the individual predicts his imperfect recall in the midmonth, is it possible to take some actions to prevent overspending in the midmonth? When he knows he will forget partial information about the past consumption in the future, will he behave differently and choose a different consumption plan at the beginning of the month? Assume the individual is able to track his consumption for the past week. Thus, the individual is able to change his consumption plan for the next week. Then, the individual on a different date could be modelled as a different self and each self is allowed to control his current one-shot behaviour.

The Brief Introduction of Chapter 1

In a decision problem with imperfect recall, an ex-ante optimal strategy is not necessarily time consistent. It indicates that the decision maker might deviate to a more beneficial strategy based on the belief consistent with the chosen ex-ante optimal strategy if he is allowed to choose any strategy at the node where the decision maker reconsiders. Then, Piccione & Rubinstein (1997a) propose the modified multiself consistency. They conclude that an ex-ante optimal strategy is modified multiself consistent. It illustrates that the decision maker will not deviate to any one-shot pure behaviour when the decision maker is allowed to reconsider once in one round play of the decision problem and he assumes his behaviours follow the ex-ante chosen strategy at the other nodes. It is interesting to see to what extent an ex-ante optimal strategy keeps its conditional optimality when the decision maker reconsiders during execution. Additionally, the nature of absentmindedness makes the decision maker reconsider more than once if the decision maker chooses where in the decision problem tree he would like to reconsider. For example, it is impossible for a decision maker to reconsider at one node in an information set with absentmindedness. When he moves to an information set with absentmindedness, he does not know he has previously reconsidered or at this information set. He must reconsider at each node in the information set along a branch of the decision problem tree. Therefore, it is interesting to explore whether a decision maker chooses to deviate from the ex-ante optimal strategy if he is allowed to reconsider one-shot deviations multiple times.

Another problematic issue about the interpretation of absentmindedness is the belief system in the conditional expected payoff function. A belief system is a collection of probability measures which indicates the probability of reaching each history conditional on reaching the information set containing those nodes. The belief system is said to be consistent with a strategy σ if it is calculated by following Bayesian rule, given the strategy σ. However, the relative probability of being at each node in an information set is determined by the strategy assigned at that information set. Then, it might cause the problematic issue of strategy-belief-strategy cycle. The cycle illustrates the fact that a decision maker's strategy at an information set determines the belief system of that information set, and the belief system reversely influences the choice of strategy. We would like to introduce the stage of planning to the procedure of decision making to solve this problem. When a decision maker reevaluates the decision problem during execution, he forms his belief system based on the strategy chosen in the stage of planning. By doing so, the processes of belief forming and strategy choosing separate. Another problem is the mutual exclusion issue. The event of reaching a node is different from the event of being at a node when absentmindedness presents. The events of reaching each node in the same information set are not necessarily mutually exclusive. The sum of the probability of reaching each node in an information set with absentmindedness might be larger than 1 and it is not the probability of reaching the information set. Thus, the probability calculated according to the definition of a consistent belief system might not be the probability of being at each node conditional on being at that information set. We attempt to solve this problem in the Appendix I.

We develop our approach in a multiself form since we analyse the situation of multiple reconsideration about one-shot deviations. A distinct self is located at each node in the decision problem. We call those selves, doers. It is identical to the modified multiself approach defined in Piccione & Rubinstein (1997a). We call the period that the doers act, the stage of execution. Besides that, an extra self is defined before the decision problem starts. A doer chooses his one-shot behavioural rule based on the conditional expected payoff function. Our model is different from the modified multiself approach in an extra self, the planner, is also defined. The planner acts before the decision problem starts. The planner chooses a strategy for the whole play according to his belief on how doers behave. If the planner believes the doers will directly follow his strategy, we say such a planner is confident. The planner is confident that the doers obey his orders. If the planner believes that the doers will reevaluate the decision problem and choose the conditionally optimal one-shot behavioural rule themselves, we say such a planner is knowledgable. The knowledgable planner learns more about doers than a confident planner. A knowledgable planner knows not only how doers behave but also the act that a doer forms belief on the probability distribution of being at each history in the same information set by his chosen strategy. The knowledgable planner knows his chosen strategy is common knowledge among doers. A planner holds different beliefs on how doers behave. In such a way, we could incorporate the framework of psychological games into our approach. In our model, we could interpret the planner as a manager of the organisation who should make a plan for the whole decision problem, and the doers as agents who could choose to follow the manager's plan or reconsider whether there is a better move.

By introducing a planner into our multiself model, our approach turns the process of decisions into a two-step procedure from a one-step procedure. For the behavioural rule at each node, the decision maker has to decide twice, once in the planning stage and once in the execution stage. It seems to make the decision problem more complicated. However, our model describes a closer situation of how individuals behave when imperfect recall presents, especially absentmindedness. In the situation of absentmindedness, neither a node the decision maker passes nor the strategy assigned at that node could be delivered when the decision maker moves to a later node in the same information set. The only information regarding his current location in the information set is the ex-ante chosen strategy, such as the second intersection in the example of absentminded driver. Besides that, it can model the decision maker's different reactions when he notices future imperfect recall from the way the decision maker acts when the decision problem has perfect recall. When a decision maker in the stage of planning knows how he will consider and behave when he meets imperfect recall, he should take the knowledge into consideration when he chooses a strategy.

In our approach, the decision maker could even deviate multiple times in the same information set if the decision problem presents absentmindedness. The agents in the same information set cannot communicate with each other, but they know the fact that the other agents might also reevaluate when they reevaluate their behaviour. Like in the paradox, the decision maker could reconsider the decision problem both at the first and second intersections. It reflects at the conditional expected payoff function of a doer that the one-shot behavioural rules of other doers are assumed to be not identical to the planner's strategy. Compared, in the modified multiself approach, the driver reassesses either at the first intersection or the second intersection. In an information set with absentmindedness, if a decision maker is assumed to reconsider once, he must be stopped by an external device. The decision maker does not know his accurate current node, not even whether he has reconsidered or not. Like in the paradox, when the driver comes to the second exit, if he knows he should reconsider in the information set, he cannot remember whether he has passed the first exit, not even whether he has reconsidered in the current information set. Then, he will also reconsider at the second intersection. It could be better understood if we explain it using the example of a manager and agents. The decision problem must be reconsidered multiple times if we allow all the agents to reevaluate their own actions. Therefore, we would like to explore if no external device can intervene in the execution of the decision problem, but instead, the decision maker would like to reevaluate his strategy at an information set (or even multiple information sets) and choose their optimal current behaviours, what the results of execution could be. We call the model discussed in this paper, psychological mutiself approach. Denote Σ a finite extensive decision problem, $PG(\Sigma)$ the psychological multiself game form of the decision problem Σ.

Then, we define a psychological multiself equilibrium with a confident planner (PMEC). The equilibrium composes a strategy from the planner, a collective strategy of doers and the planner's first order belief on doers' strategies. In the definition of a psychological equilibrium, the player's belief on other players' strategies should be identical to players' equilibrium strategies. Thus, a PMEC requires that, in equilibrium, the planner's strategy, the collective strategy of doers and the planner's belief are identical to each other. Besides that, the doers cannot deviate to another equilibrium among doers since we consider multiple reconsiderations in the model. When a doer reevaluates the decision problem, he assumes the one-shot behaviours at other nodes can be arbitrary, which indicates the other doers could deviate from the planner's strategy. If the collective strategy of doers in equilibrium is the same as the equilibrium strategy, there is not a group of doers such that each of them deviates to a different one-shot behavioural rule from the planner's equilibrium strategy. Then, it implies that, for any doer, other equilibrium among doers cannot bring him strictly higher expected payoffs. Assume that the doer will not deviate if it is not strictly beneficial for him to change his current behaviour. Therefore, a strategy from a PMEC is a stable equilibrium strategy among doers. There might be asymmetry equilibrium among doers such that the equilibrium strategies of doers in the same information set are not identical. However, we assume the internal consistency of a decision maker. Internal consistency could be understood as the way that a decision maker assumes the same behaviours at other nodes if he faces an identical situation. Additionally, the decision maker should behave the same in equilibrium if he is not able to distinguish different situations. Then, in equilibrium, the strategies of the doers in the same information set should be identical. We prove that a strategy is ex-ante optimal in the decision problem Σ if and only if it belongs to a PMEC in psychological multiself game form of the decision problem $PG(\Sigma)$.

The concept of multiple multiselves consistency is defined. A strategy is multiple multiselves consistent, if reaching any node, the decision maker will not deviate even if he is allowed to reconsider multiple times. The definition of multiple multiselves consistency is almost the same as that of a PMEC except that a strategy is multiple multiselves consistent if it is conditionally optimal among all behavioural strategies in Σ while a strategy is from a PMEC indicates it is conditionally optimal among all collective strategy of doers. Then, an ex-ante optimal strategy is multiple multiselves consistent. Therefore, the decision maker would not deviate from an ex-ante optimal strategy even if the decision maker is allowed to reconsider a one-shot deviation multiple times. It requires that if a decision maker reconsiders at a node, he must reconsider at every node in the same information set as the node since the available deviation is a behavioural strategy of Σ. Assume the deviation in the form of a behavioural strategy exists at an information set X, the deviation cannot be realised if the decision maker does not reconsider at one node in that information set.

We also define a psychological multiself equilibrium with a knowledgable planner (PMEK). The equilibrium composes a strategy from the planner, a collective strategy of doers and the planner's first order belief on doers' strategies. We do not expect the equilibrium strategy of a planner and the equilibrium collective strategy of doers to be the same in a PMEK. However, in equilibrium, the planner's belief should be identical to the equilibrium collective strategy of doers. It is the reason why we interpret a decision problem as a two-step procedure. It presents how the decision maker actually behaves if he notices the future imperfect recall (absentmindedness) and changes the way he evaluates the decision problem which makes it different from the situation of perfect recall.

The Brief Introduction of Chapter 2

In classic decision problem theory, there is not much difference between the following two situations when a decision maker reevaluates a finite decision problem with imperfect recall. The first one is that the decision maker knows he will present imperfect recall at a later information set. The second situation is that the decision maker does not know he will present imperfect recall but he is required to choose the same behavioural rule at those nodes which are in the same information set. Then, the assumption that the decision maker knows the structure of the decision problem tree contributes nothing in the process of decision making. In other words, it seems to be indifferent whether a decision maker knows he will present imperfect recall as the decision problem executes. It is interesting to explore how a decision maker would behave if he prepares for his future imperfect recall when he reconsiders during execution. We call such a decision maker sophisticated.

In a finite decision problem with perfect recall, the conditional expected payoff function could be written as a recursive form such that, for an information set X and all information sets X_j which are reached by one-shot move from a node in the information set X, the conditional expected payoffs at X is the sum of the probability of reaching X_j from X multiplies the conditional expected payoffs at X_j. The conditional expected payoff function at an information set X contains the information on how a decision maker considers the decision problem. Then, we wonder whether it is possible to develop a recursive calculation rule to model how a sophisticated decision maker reconsiders the decision problem with imperfect recall during execution.

A sophisticated decision maker at an information set explains his current one-shot behaviour as it leads him to another information set instead of a history (the one-shot behaviour could also lead him to the same information set if the current information set where he is located presents absentmindedness). The attitude modified IS-path probability function in the sophisticated recursive calculation rule describes the situation that how a decision maker evaluates the decision problem at an information set X. Firstly, he forms a belief system consistent with the strategy σ he is executing now. The belief system indicates the probability of being at a node $h \in X$ conditional on the information set X. Conditional on being at the node h, he could actually implement the action a. The probability that he implements the action a is described by the strategy σ. Once the decision maker implements the action a, it could lead him to either the same information set if information set presents absentmindedness or another information set. If the action leads him to the same information set, the decision maker expects to obtain the full conditional expected payoffs. If the action leads him to another information set, the decision maker knows it is possible that he could also reach that information set through a path which does not contain any node from the information set X. Then, the decision maker expects λ of the conditional expected payoffs at that information set. The parameter λ conveys the viewpoint that the decision maker at the beginning expects to obtain identical expected payoffs of an information set no matter which path he passes through. By this step, the attitude modified IS-path probability function calculates the attitude modified probability from the information set X to each information set that the decision maker can reach with one-shot move.

If we repeat this step at those information sets which can be reached by the decision maker with one-shot move, it calculates the probabilities to reach other new information sets. If we keep repeating this step until a terminal node. By recursion, it calculates the attitude modified probability of an IS-path from the information set X to any terminal node. We say there is an IS-path from information set A to information set B if, for any adjacent two information sets in the IS-path, the decision maker could reach one of those two information sets to the other one by one-shot move. Then, there is another problem that a terminal node could be considered in more than one IS-path. However, there is at most one IS-path containing the real path from the decision maker's current information set to any terminal node. Some terminal nodes that will never be reached conditional on the current information set should also be considered currently. It is because the decision maker knows the structure of the decision problem and the fact he will lose the information of the current history he is located if the information set reached by his one-shot move presents imperfect recall. The terminal nodes considered at those information sets he thinks he might reach in the future due to imperfect recall should also be considered now.

The average function solves the problem of multiple IS-paths from an information set to a terminal node. Since the decision maker act as if he cannot figure out which IS-path contains the real path to a terminal node when there is more than one IS-path, the decision maker is assumed to assign equal probability to each one. Then, we apply the attitude modified IS-path probability function to the substitute function. We call the resulted conditional expected payoff function IS expected payoff function. The IS expected payoff function implies the way a sophisticated decision maker considers conditional on an information set. We are going to prove that a strategy is ex-ante optimal if and only if it is IS-time consistent for a finite decision problem without crossing information sets. Two information sets X_1 and X_2 are defined to be crossing information sets, if $\exists h_1, h_1' \in X_1$, $\exists h_2, h_2' \in X_2$, such that h_1 is a subhistory of h_2, and h_2' is a subhistory of h_1'. A strategy is IS-time consistent if it is conditionally optimal at any information set if the decision maker evaluates by the IS expected payoff function (there are some requirements for an available deviation strategy at an information set, see section 6 of chapter 2). It implies that the decision maker cannot find a more beneficial strategy than the optimal strategy he chooses before the decision problem starts if he prepares for his future imperfect recall when he evaluates the decision problem at any information set. In other words, in the middle of a decision problem, an individual never denies the best plan he chooses ex-ante if he always takes how he will evaluate conditional on each future step into consideration when he evaluates now.

The Brief Introduction of Chapter 3

In this chapter, we apply the sophisticated recursive calculation rule to a finite game with imperfect recall. For simplicity, the game is assumed to have complete information and the other players have perfect information except for one player who has imperfect recall. Even for the player who has imperfect recall, we assume he has perfect information about the other players' strategies. In fact, it does not matter whether there is more than one player who presents imperfect recall as long as the players have perfect information about the other players' strategies. The imperfect recall is an internal imperfection. The third party cannot observe it directly but might deduce it through the player's “irrational" behaviours. If we apply the sophisticated recursive calculation rule to the finite game, the conditional expected payoff function for the player with imperfect recall at an information set becomes IS expected payoff function.

A sophisticated player knows he will present imperfect recall in the future execution of the game. At that information set with imperfect recall, the player does not know which node he is currently located. The player has to choose a strategy based on the belief system he forms which is consistent with the strategies of the other players and the strategy he executes until the current node. Then, he might consider some outcomes which will never be achieved by any strategy due to his imperfect recall. Although the player at the current node might be able to know which outcomes are impossible to realise at that information set with imperfect recall, he would rather behave as if he does not know the information about the outcomes. Then, it is possible for him to choose a strategy at the current node which is also conditionally optimal when he reconsiders at that information set since he currently considers the game in the same way as if he moved to the future information set. We would like to examine the sequential rationality of a strategy. In this chapter, we use IS expected payoff function as the standard to examine the sequential rationality of a strategy.

The main motivation is to see whether the sophisticated recursive calculation rule can be applied to a finite game. Furthermore, if the resulted IS expected payoff function is consistent with that in a finite decision problem, it is interesting to see whether a strategy is ex-ante optimal if it is sequentially optimal, given the strategy profile of the other players.

The IS expected payoff function at the information set which contains the first node where a player acts coincides with the ex-ante expected payoff function in conventional game theory. Thus, an IS sequential equilibrium must be a Nash equilibrium. The existence of an IS sequential equilibrium indicates that there exists a strategy profile such that, conditional on the perfect knowledge about the other players, no player has a conditionally more beneficial strategy than the strategy described in the equilibrium strategy profile when any player reconsiders at any information set, if the player with imperfect recall is sophisticated.

Besides the IS sequential equilibrium, we also discuss the inequivalence between behavioural strategies and mixed strategies in non-linear finite games, such as absentminded driver paradox. Therefore, we have to introduce the concept of general strategies. A general strategy involves two-stage randomisation over pure strategies: ex-ante and during execution. Therefore, it includes both behavioural strategies and mixed strategies. The analysis of imperfect recall in a finite game should be discussed over general strategies.

What we have done in the dissertation is to find different frameworks to interpret imperfect recall. Including absentmindedness causes much more difficulty to develop a reasonable model. In the first chapter, we choose to analyse a finite decision problem with imperfect recall in a multiself approach. It is motivated by the situation of absentmindedness. A decision maker at an information set presenting absentmindedness cannot convey the information of which nodes he has passed and the strategies chosen at those nodes. Then, the most reasonable assumption for the reconsideration of a decision maker at an information set with absentmindedness is to have one-shot deviation. Then, it is natural to model the decision problem as if a distinct self moves at each node.

The multiself approach transfers the decision problem into several smaller decision problems which analyse the decision maker more meticulously. Then, it is interesting to see that what if we merge the situations that a decision maker is confused into one object and solve the decision problem coarsely. The sophisticated decision maker (player) defined in chapter 2 (3) presents the viewpoint in the way that he does not care what happens within an information set except his current one, what the decision maker knows is the directed connection among information sets. The decision maker gives up to confirm his accurate history where he is located in an information set with imperfect recall from the beginning of the decision problem. It is interesting to explore whether an ex-ante optimal strategy is conditionally optimal under such assumptions. The decision maker (player) in this model is assumed to control his current behaviour and the behaviours at those information sets which can be reached if he moves out of the current one. Although we assume the decision maker (player) is able to control his current behavioural in the current information set, we also propose that the decision maker believes he would behave the same if he also reconsiders at the other nodes in the same information set. It could be explained by the internal consistency of an individual that he chooses the same way if he faces the same decision problem successively.

Chapter 1

Psychological Multiself Approach: How Does a Decision Maker with Absentmindedness React to Time Inconsistency?

1.1 Introduction

In an extensive decision problem, without any changes in preferences and any new information about the decision problem, we say a strategy is time consistent if there is not, at any information set that could be reached by implementing the strategy, a different strategy yields a higher conditional expected payoff. A decision problem is defined to have perfect recall if the decision maker remembers all the information previously acquired and all his past actions. In decision problems with perfect recall, a strategy is ex-ante optimal if and only if it is time consistent. Then, it does not matter whether a decision maker makes all the decisions before the decision problem starts or the decision maker makes the current decision whenever he moves to a new situation. The analysis of equilibrium issues is much easier if the decision problem presents perfect recall.

However, in real life, most individuals have imperfect recall. Sometimes, it is hard even if we are required to remember all of our past actions in a game, such as Die-roll poker. Another example could be the online banking or the use of credit cards. If an individual makes the consumption plan for the next month, he might overspend due to his imperfect memory of the use of credit card. In this paper, we also discuss a manager-agents model, in which the manager and agents share the same payoff function but lack communication among agents. Once the assumption of perfect recall is removed, an ex-ante optimal strategy may not be time consistent, and vice versa. An ex-ante optimal strategy maximises the expected payoffs when the decision maker evaluates the decision problem before it plays. The inequivalence of ex-ante optimality and time consistency arises paradoxes in the viewpoint of standard decision theory. Besides it, there are many ambiguities in the interpretation of decision problems with imperfect recall, for example, conditional expected payoffs on an information set that exhibits imperfect recall. Thus, since Kuhn (1953) introduces the concepts of perfect and imperfect recall, most recent discussions over extensive decision problems and games are confined to those with perfect recall.

We would like to illustrate the inequivalence between ex-ante optimality and time consistency, as well as some ambiguities in the interpretation of decision problems with imperfect recall by a famous example, absentminded driver paradox. The absentminded driver paradox is proposed by Piccione & Rubinstein (1997a) (PR from now on). It concretises a special case of decision problems with imperfect recall, absentmindedness. Absentmindedness refers to a situation in which the decision maker is not able to distinguish between two or more histories along the same decision tree path. It illustrates that the decision maker might even forget the fact that he has made a decision previously. The paradox is described as follows (also see Figure 3.1).

> *A drunk person sits in a bar at night considering his midnight trip home. In order to arrive home, he needs to take the highway and get off at the second intersection (payoff 4). However, tuning at the first intersection will lead to a disastrous place (payoff 0). Once he continues on the highway beyond the second intersection, he can never go back tonight and need to find a hotel where he can spend the night there (payoff 1). Because of being drunk, this person knows that he will not be able to remember how many intersections he has passed by. Besides, there is no external device to help the person know his current location and it is impossible to figure out by observing the around areas at any intersection.*

We start with pure strategies. When planning his journey in the bar, the drunk person should know he will never arrive home. Before driving onto the highway, he needs to choose to either continue (action C) or exit (action E) at each intersection. The actions at both intersections should be identical since he knows when reaching an intersection, he is not able to recognise which one it is. Thus, he will exit at the first intersection e_1 if he chooses to exit at the bar and obtain payoff 0. He will end up sleeping in a hotel if he chooses to continue at both intersections, and obtain payoff 1. The ex-ante optimal pure strategy is obviously to continue at both intersections which helps him obtain payoff 1. However, once reaching one intersection, the decision maker would deduce he is at the first intersection if he had chosen to exit ex-ante. His best response is to continue instead of following his strategy to exit, the former yields a positive expected payoff which is higher than the latter choice (payoff 0). Similarly, the decision maker would assign $1/2$ to being at each intersection if he chooses to continue as his strategy in the bar. When the decision maker executes the ex-ante strategy that he should continue at both intersections. The probability of reaching the first intersection is 1 and the probability of reaching the second intersection is 1. Apply Bayesian rule to calculate the relative probability of being at each intersection. The probability of being at the first intersection and the second intersection is equal, $1/2$. Thus, a decision maker with ex-ante pure strategy of continuing at both intersections assigns equal probability to each intersection. At this point, he should change from action C (payoff 1) to E (expected payoff 2). Therefore, the decision maker would like to change his plan no matter what his initial strategy is although there is no new information or changes in his preferences.

When the decision calculates the relative probability of being at each intersection in this example, he ignores the fact of mutual exclusion and the difference between being at an intersection and passing an intersection. Assume the driver implements the pure strategy of continuing at both intersections. For him, the probability of being (passing) the first intersection is 1 and the probability of being (passing) the second intersection is 1. Then, he thinks the sum of the probability of being at either the first intersection or the second intersection is 2. Then, the relative probability of being at the first intersection conditional on being at either the first or second intersection is $1/2$. In the same way, the relative probability of being at the second intersection conditional on being at either the first or second intersection is also $1/2$. However, the way to calculate the relative probability of being at a node is not true when absentmindedness presents. The probability of passing either the first intersection or the second intersection is 1 instead of 2 since the driver must have passed the first intersection if he passes the second intersection. Besides it, passing a node can be understood as that the decision maker has passed the node after the decision problem ends. It is the ex-post probability. Being at a node means the decision maker is located at the node when the decision maker is interrupted in the middle of the play. Assume a decision maker passes the second intersection. It includes two possibilities. The first one is that the decision maker is located at the first intersection when the decision problem is stopped. The second one is the decision maker is located at the second intersection when the decision problem is stopped. It is not true to use the probability of passing the second intersection to calculate the probability of being at the second intersection.

Although the above way is not reasonable, there is not a convincing way to calculate the probability of being at a node if the decision problem is stopped during execution. Most works including our results continue to use the rule of consistent belief as the way to calculate the probability of being at a node. We provide a sketch to define a reasonable belief system in Appendix I.

The discussions on time consistency allow a decision maker reassesses his strategy at any history during execution. It focuses on the procedural aspects of decision problems. At a history where the decision maker stops and reconsiders the decision problem, he has two ways of reasoning. One way of reasoning is, once having chosen a strategy before the play starts, the decision maker should stick to the determined strategy without verifying its optimality unless there is a change in his preference or new information about the decision problem intervenes. He believes that he has deliberated and chosen the best strategy for himself ex ante. The other one is to adjust his following behavioural plan to make it maximise the expected payoffs conditionally at his current history. Under the assumption of perfect recall, the decision maker has no other better choice but to follow the ex-ante chosen optimal strategy throughout the execution, thus, two ways of reasoning reach the same result if he executes an optimal strategy. However, when imperfect recall presents, an optimal strategy could be time inconsistent. The decision maker might have a better different action plan which obtains higher conditional expected payoffs than the ex-ante optimal strategy at some history. He needs to choose between following the original strategy or changing in a new plan, which works out different results. It is hard to judge which way of reasoning shows an individual's rationality. The reason why the absentminded driver example makes a paradoxical scenario is that it demonstrates two conflicting reasoning as above. In that case, the timing that a decision maker chooses his behaviours matters if one would like to predict the results of a decision problem.

Since the paradox appears no matter what pure strategy is adopted, it is naturally to consider whether taking randomisation into consideration can solve the paradox. Denote the probability to continue is p at an intersection. Before the person starts onto the highway (before the decision problem starts), his expected payoff is

$$p^2+4p(1-p) = -3p^2+4p = -3(p-2/3)^2+4/3.$$

Clearly, $(p,p) = (2/3,2/3)$ is the optimal strategy for this example. However, when he comes to some intersection, holding the belief of α to being at the first cross, his expected payoff becomes

$$\begin{aligned}\alpha(p^2+4p(1-p))+(1-\alpha)(p+4(1-p)) &= -3\alpha p^2+(7\alpha-3)p-4\alpha+4\\ &= -3\alpha(p-\frac{7\alpha-3}{6\alpha})^2+\frac{\alpha^2-42\alpha+57}{12\alpha}.\end{aligned}$$

The strategy which maximises the conditional expected payoff becomes $p = \max\{0, (7\alpha-3)/6\alpha\}$, which is different from the ex-ante optimal strategy $(p,p) = (2/3,2/3)$ unless $\alpha = 1$. That is to say, the optimal strategy is not time consistent unless the decision maker believes he is certainly at intersection e_1 when he is going to make a decision in the information set I. However, the decision maker has no confidence to confirm that he must be at the first intersection when he realises he is on the highway. A more reasonable belief should be $\alpha = 3/5$ with respect to the strategy $(p,p) = (2/3,2/3)$. It follows Bayes' rule in notation, which is identical to the calculation rule of consistent belief defined in games with imperfect information in classic game theory. In general, the reasonable belief is $\alpha = 1/(p+1)$ with respect to the strategy (p,p). With this belief, the optimal strategy $(p,p) = (2/3,2/3)$ is not time consistent. Therefore, introducing randomisation cannot help vanish the paradox.

Which strategy is time consistent in this example? To calculate it, denote (p,p) is the strategy whose optimality is ready to be checked and (q,q) is the strategy that the decision maker would like to choose when he reevaluates the expected payoffs in the midst of the decision problem. During the execution, his expected payoffs become

$$EU(q|p) = \frac{1}{p+1}(q^2+4q(1-q))+\frac{p}{p+1}(q+4(1-q)) = \frac{1}{p+1}(-3q^2+(4-3p)q+4p).$$

The decision maker should choose q with a fixed known p.

$$q^{\star} = \underset{q}{argmax} \, EU(q|p^{\star}) = p^{\star}$$

is required if $(p^{\star}, p^{\star})$ is the time consistent strategy. In conclusion, $(p^{\star}, p^{\star}) = (4/9, 4/9)$ is the unique time consistent strategy for this example. Therefore, in this example, the ex-ante optimal strategy is not time consistent, in the meanwhile, the time consistent strategy is not ex-ante optimal strategy.

In PR's paper, they propose not only the classic paradox to explain absentmindedness but also the modified multiself approach to solve the problem caused by absentmindedness. The approach states that, once a decision maker is stopped in the middle of execution, instead of persisting that he is able to control all of his following behaviours, PR manifest the decision maker's control power could be restricted to his current move. Thus, we could regard a decision maker at different nodes in the decision problem trees as different selves. Then, PR expound that an ex-ante optimal strategy is modified multiself consistent. In other words, an ex-ante strategy is time consistent if, at most we allow the decision maker deviates to a strategy which is identical to the initial one except his current behaviour. We call such a strategy has **one shot deviation** from the original strategy if the strategy assigns the same behavioural rule to every node as the original strategy except one node. Normally, imperfect recall is understood as an individual who has limited ability of memory. However, in this paper, we would rather explain imperfect recall as a team with communication issues. As stated in PR, the decision maker works like an organisation consisting of agents who act once and successively. All the agents have identical payoffs. The collection of agents' behaviours is equivalent to a strategy of the decision maker.

In this paper, we would like to extend PR's multiself approach in the following aspects. Firstly, a distinct self is defined at each history in the identical way to that defined in modified multiself approach. We call those selves doers and a doer acts when the decision maker moves to the node where the doer is located during execution. We call the period that the doers act, the stage of execution. A doer chooses his one-shot behavioural rule based on the conditional expected payoff function. Our model is different from the modified multiself approach in an extra self, the planner, is also defined. The planner acts before the decision problem starts. The planner chooses a strategy for the whole play according to his belief on how doers behave. If the planner believes the doers will directly follow his strategy, we say such a planner is confident. The planner is confident that the doers obey his orders. If the planner believes that the doers will reevaluate the decision problem and choose the conditionally optimal one-shot behavioural rule themselves, we say such a planner is knowledgable. The knowledgable planner learns more about doers than a confident planner. A knowledgable planner knows not only how doers behave but also the act that a doer forms belief on the probability distribution of being at each history in the same information set by his chosen strategy. The knowledgable planner knows his chosen strategy is a common knowledge among doers. A planner holds different beliefs on how doers behave. In such a way, we could incorporate the framework of psychological games into our approach. In our model, we could interpret the planner as a manager of the organisation who should make a plan for the whole decision problem, and the doers as agents who could choose to follow the manager's plan or reconsider whether there is a better move.

By adding a planner into our multiself model, our approach turns the process of decisions into a two-step procedure from a one-step procedure. For the behavioural rule at each node, the decision maker has to decide twice, once in the planning stage and once in the execution stage. It seems to make the decision problem more complicated. However, our model describes a closer situation of how individuals behave when imperfect recall presents, especially absentmindedness. In the situation of absentmindedness, neither a node the decision maker passes nor the strategy assigned at that node could be delivered when the decision maker moves to a later node in the same information set. The only information regarding his current location in the information set is the ex-ante chosen strategy, such as the second intersection in the example of absentminded driver. Besides that, it can model the decision maker's different reactions when he notices future imperfect recall from the way the decision maker acts when the decision problem has perfect recall. When a decision maker in the stage of planning knows how he will consider and behave when he meets imperfect recall, he should take the knowledge into consideration when he chooses a strategy.

Different from the modified multiself approach whose underlying assumption is that a decision maker could reconsider once in an occurrence of a decision problem, we assume that a decision maker would like to deviate more than once. The decision maker could even deviate multiple times in the same information set if the decision problem presents absentmindedness. The agents in the same information set cannot communicate with each other, but they know the fact that the other agents might also reevaluate when they reevaluate their behaviour. Like in the paradox, the decision maker could reconsider the decision problem both at the first and second intersections. It reflects at the conditional expected payoff function of a doer that the one-shot behavioural rules of other doers are assumed to be not identical to the planner's strategy. Compared, in the modified multiself approach, the driver reassesses either at the first intersection or the second intersection. In an information set with absentmindedness, if a decision maker is assumed to reconsider once, he must be stopped by an external device. The decision maker does not know his accurate current node, not even whether he has reconsidered or not. Like in the paradox, when the driver comes to the second exit, if he knows he should reconsider in the information set, he cannot remember whether he has passed the first exit, not even whether he has reconsidered in the current information set. Then, he will also reconsider at the second intersection. It could be better understood if we explain it using the example of a manager and agents. The decision problem must be reconsidered multiple times if we allow all the agents to reevaluate their own actions. Therefore, we would like to explore if no external device can intervene in the execution of the decision problem, but instead, the decision maker would like to reevaluate his strategy at an information set (or even multiple information sets) and to choose their optimal current behaviours, what the results of execution could be. We call the model discussed in this paper, psychological mutiself approach. Denote Σ a finite extensive decision problem, $PG(\Sigma)$ the psychological multiself game form of the decision problem Σ.

Then, we define a psychological multiself equilibrium with a confident planner (PMEC). The equilibrium composes a strategy from the planner, a collective strategy of doers and the planner's first order belief on doers' strategies. In the definition of a psychological equilibrium, the player's belief on other players' strategies should be identical to players' equilibrium strategies. Thus, a PMEC requires that, in equilibrium, the planner's strategy, the collective strategy of doers and the planner's belief are identical to each other. Besides that, the doers cannot deviate to another equilibrium among doers since we consider multiple reconsiderations in the model. When a doer reevaluates the decision problem, he assumes the one-shot behaviours at other nodes can be arbitrary, which indicates the other doers could deviate from the planner's strategy. If the collective strategy of doers in equilibrium is the same as the equilibrium strategy, there is not a group of doers such that each of them deviates to a different one-shot behavioural rule from the planner's equilibrium strategy. Then, it implies that, for any doer, other equilibrium among doers cannot bring him strictly higher expected payoffs. Assume that the doer will not deviate if it is not strictly beneficial for him to change his current behaviour. Therefore, a strategy from a PMEC is a stable equilibrium strategy among doers. There might be asymmetry equilibrium among doers such that the equilibrium strategies of doers in the same information set are not identical. However, we assume the internal consistency of a decision maker. The internal consistency could be understood as the way that a decision maker assumes the same behaviours at other nodes if he faces an identical situation. Additionally, the decision maker should behave the same in equilibrium if he is not able to distinguish different situations. Then, in equilibrium, the strategies of the doers in the same information set should be identical. We prove that a strategy is ex-ante optimal in the decision problem Σ if and only if it belongs to a PMEC in psychological multiself game form of the decision problem $PG(\Sigma)$.

The concept of multiple multiselves consistency is developed. A strategy is multiple multiselves consistent, if reaching any node, the decision maker will not deviate even if he is allowed to reconsider multiple times. The definition of multiple multiselves consistency is almost the same as that of a PMEC except that a strategy is multiple multiselves consistent if it is conditionally optimal among all behavioural strategies in Σ while a strategy is from a PMEC indicates it is conditionally optimal among all collective strategy of doers. Then, an ex-ante optimal strategy is multiple multiselves consistent. Therefore, the decision maker would not deviate from an ex-ante optimal strategy even if the decision maker is allowed to reconsider a one-shot deviation multiple times, with the requirement that if a decision maker reconsider at a node, he must reconsider at every node in the same information set as the node since the available deviation is a behavioural strategy of Σ. Assume the deviation in the form of a behavioural strategy exists at an information set X, the deviation cannot be realised if the decision maker does not reconsider at one node in that information set.

We also define a psychological multiself equilibrium with a knowledgable planner (PMEK). The equilibrium composes a strategy from the planner, a collective strategy of doers and the planner's first order belief on doers' strategies. We do not expect the equilibrium strategy of a planner and the equilibrium collective strategy of doers to be the same in a PMEK. However, in equilibrium, the planner's belief should be identical to the equilibrium collective strategy of doers. It is the reason why we interpret a decision problem as a two-step procedure. It presents how the decision maker actually behaves if he notices the future imperfect recall (absentmindedness) and changes the way he evaluates the decision problem which makes it different from the situation of perfect recall.

The paper is organised as follows. The second section introduces previous research on absentmindedness. The third section illustrates an example of absentmindedness in detail. In the fourth section, we present past works about psychological games and explain why we incorporate psychological games into our approach. Then, we apply our approach to the example of absentminded driver paradox to have an intuitive understanding of a PMEC and a PMEK. In the next section, the concepts of a psychological multiself game form of a decision problem, a PMEC, a PMEK and multiple modified consistency are defined and relevant propositions are discussed. The last part of our paper is the conclusion and extension.

1.2 Previous Works of Decision Problems with Absentmindedness

There are a bunch of papers which propose different models to explain the time inconsistency problem in the example of the absentminded driver paradox, even in a general finite decision problem with imperfect recall. They can be divided into two categories, multiself approaches and one-self approaches.

Let us start with the comparison of multiself approach and one-self approach. Strotz (1955) is the first one who provides a multiself perspective to explain a decision problem (game). It is a framework in which every information set is assumed to be a point of decision and the decision maker at each decision point is unable to control his behaviour at future information sets. A decision maker works as a collection of selves with identical preference whose behaviours form an equilibrium. Without absentmindedness, the decision maker acts once at any information set in an occurrence of the play. Multiself approach restricts the decision maker's control power to his current behaviour. We can see that the difference between one-self and multiself approach is the control power of the decision maker. The two approaches show extreme assumptions. The decision maker is able to control full of his future actions in a one-self approach while he can only change his current decision in a multiself approach. In general, the requirement of no deviation in the one-self approach is stronger than that in the multiself approach. However, with the presence of perfect recall, an ex-ante optimal strategy is an equilibrium of selves for the game version of a decision problem, and vice versa. Combined with the fact that a strategy is ex-ante optimal if and only if it is time consistent, it implies the decision maker who executes an ex-ante optimal strategy has no motivation to deviate from the behavioural rule assigned to his current move when reaching an information set. It is also called "no single improvement" property. In other words, no single improvement is equivalent to time consistency under the condition of perfect recall. In fact, the two properties (time consistency and no single improvement) are the measurement of conditional optimality of a strategy during execution in one-self approach and multiself approach, respectively.

The absentminded driver paradox from PR sheds light on the issues of absentmindedness in decision problems. When an information set presents absentmindedness, the decision maker who passes it might move more than once in that information set during one round play of the decision problem. The “no single improvement" property in perfect recall is not enough to describe the decision maker’s control power in absentmindedness. No single improvement property does not clarify the decision maker can control his behaviour over the current move or the current information set since they are equivalent in the presence of perfect recall. However, with absentmindedness, the two assumptions illustrate opposite situations. The decision maker will follow the initial strategy when the same information set is revisited if he is assumed to control his immediate behaviour while the decision maker will follow whatever behavioural rule he adopts now when he visits the same information set again if the decision maker is assumed to control his behaviours at the current information set. The modified multiself approach in PR discusses decision problems with imperfect recall under the first assumption that the decision maker is able to control only his current action while assuming other visits of the same information set follow the original strategy. The “modified" multiself approach modifies the multiself analysis by Strotz (1955) that every self of the decision maker is located at each history instead of information set since there is more than one history along a decision tree path if absentmindedness presents. The modified multislf approach explains that time inconsistency vanishes if the decision maker is restricted to controlling his immediate behaviour during execution.

Besides PR, Aumann et al. (1997a), Gilboa (1997) and Rabinowicz (2003) also explain the decision problems with imperfect recall in a multiself approach. Rabinowicz (2003) develops a similar framework to Aumann et al. (1997a) but the details of the argument are different. Thus, we place emphasis on the former two papers. Aumann et al. (1997a) describe the paradox as a two-stage problem, planning stage and action stage. In the planning stage, the decision maker finds the optimal strategy by maximising expected payoff; while in the action stage, the decision maker maximises conditional expected payoff with the assumption he will develop his belief from his action at other intersection and implement the behaviour in the equilibrium at other intersection. Gilboa (1997) formulates two symmetric agents of the decision maker, and nature would decide whether each agent will be called upon to act at the first intersection with equal probabilities, in which the equal probability is assigned by nature. When anyone of the two agents makes a decision, he takes the other's behaviour as given. In this way, he interprets a decision problem with absentmindedness into a two-agent game with perfect recall. The common point of these two approaches is that, during execution, the decision maker could control the behaviour at the immediate move while regarding his behaviour at other nodes as given. They are different in their attitude towards the pre-play stage. In Aumann et al. (1997a), the decision maker at other nodes follows the strategy he proposes at the planning stage, while in Gilboa (1997), the decision maker has no control beyond the instances where he moves. Lipman (1997) illustrates that the approaches in Aumann et al. (1997a) and Gilboa (1997) are equivalent in terms of the resulting equilibrium. In fact, we find the substantial reason for equivalence is that both approaches assume the equivalent way to calculate a reasonable belief system at a non-singleton information set. In Aumann et al. (1997a), nature assigns equal probability to the instance that the decision maker is stopped and reconsiders at each intersection, while in Gilboa (1997), nature assigns equal probability to the situation that the one who is reconsidering is the first active agent and the second agent respectively. From the positive aspect, the two approaches avoid the problematic cycle of "strategy-belief-strategy"(Gilboa 1997). The cycle of "strategy-belief-strategy" means that the decision maker's strategy determines his consistent belief and his belief affects the strategy he is going to choose if absentmindedness is analysed in the conventional framework of game theory. However, the resulting equilibrium is not necessarily equivalent to each other if we assume nature assigns an arbitrary probability to two scenarios in each approach.

For the one-self approach, Segal (2000), Dimitri (1999), Dimitri (2009) and Battigalli (1997) use different models to explain why the optimal strategy $(p, p) = (2/3, 2/3)$ yields the highest conditional expected payoff during the play of absentminded driver example. Segal (2000) presents the paradox as a lottery. He explains the reason why the example is paradoxical is that the ex-ante lottery should not be the same as the lottery which the decision maker faces at the information set but is defined identically in PR's paper. Dimitri 's approach has two basic assumptions, welfare symmetry and belief consistency (Dimitri 1999, 2009). Welfare symmetry indicates the decision maker values the same expected payoff at both intersections. He assumes the evaluation of the decision problem at both intersections is the same and establishes an equation to solve the expression of the evaluation and maximise it to obtain the optimal strategy. Strictly speaking, the concept of constrained time consistency in Battigalli (1997) does not belong to one-self approaches. One-self analysis assumes the decision maker has full control of his future behaviours even at nodes he might not reach while constrained time consistency guarantees a strategy is conditionally optimal if the decision maker can only change his behaviours at nodes which are not reachable by implementing the strategy.

Other issues about the interpretation of imperfect recall are also addressed by some papers. Grove & Halpern (1997) and Halpern (1997) provide new insight into the calculation of conditional expected payoffs. They explain that a decision maker only reevaluates when he reaches an information set at the first time. However, we doubt the reasonability of this assumption. Like in the example of the absentminded driver paradox, how does the driver deduce he is at the first intersection while he is reconsidering? The only possibility is the driver knows the external information that he should be asked to reassess only if he is at the upper counter of an information set. However, under that circumstance, the decision maker can definitely know it is the first time that he passes the information set. Accordingly, the driver knows he is at the first intersection when he is ready to reconsider. It is contracted to the definition of an information set in which a decision maker cannot distinguish among nodes. Besides it, Grove & Halpern (1997) and Halpern (1997) claim that it might be not enough to directly apply the framework of classic decision theory to decision problems with imperfect recall. Halpern (1997) adopts some elements used in computer sciences which incorporate statements about the track of the decision maker's knowledge of his past strategies.

1.3 Manager-agents Model

Most concepts regarding consistency, such as time consistency in conventional game theory, modified multiself consistency in PR, constrained time consistency in Battigalli (1997), gt time consistency in Halpern (1997), assumes the decision maker only reconsiders once during one round of play. For example, the modified multiself consistency in PR assumes that the decision maker implements the original strategy unless he is at the current node where he reevaluates. It illustrates that the decision maker reconsiders only at the current node during this round of execution. However, we might question whether it is a reasonable assumption that a decision maker reevaluates a decision problem at any one node during once implementation. In fact, Piccione and Rubinstein have put forward the question in their reply to responses from numbers of economists in PR (Piccione & Rubinstein 1997b). In the end of their reply, one of the questions they raise is what makes a decision maker believe that his current change of strategy is just a one shot deviation, even though he knows he will meet the identical situation later during the execution of a decision problem. Then, it should be worth exploring whether the consistency exists and what could occur if the decision maker is allowed to reconsider more than once along one play, especially when absentmindedness appears. It is our motivation to investigate a model allowing multiple times of reconsiderations.

Since our model studies multiple times of reevaluations during execution, the commitment to implement a strategy completely seems to be more challenging. Compared with the situation of reconsideration at a unique node, the decision maker now has more chances to change his initial strategy, let alone if the decision maker reassesses at each node. Therefore, it should be better accorded with the circumstance we describe if the decision maker is assumed to control his immediate action whenever reevaluating the decision problem. Furthermore, the decision maker knows he cannot remember his new strategy for the rest of play when he is at an information set featuring absentmindedness. Like in the example of the absentminded driver paradox, the driver cannot remember the strategy he deviated at the first intersection when he moves to the second intersection. The decision maker might rather choose a most beneficial current action than expect he could control his behaviours at later instances. Another reason that we use multiself approach to analyse decision problems with imperfect recall is that most individuals are short-sighted nowadays, especially when they lack precise information. The decision maker focuses on the optimal immediate behaviour instead of a long complete plan.

In our approach, we say the decision maker who chooses a strategy before the decision problem executes is a **planner**. The planner cannot make any actual moves, but rather provides a reminder or guideline for the decision maker such that the decision maker knows what he initially would like to do or is ordered to do. It works as the “basement" strategy that appears in the various definitions regarding consistency. In other words, the planner is a passive player. The decision maker who is executing a strategy is called a **doer**. The doers are located at different nodes. One distinct doer is at each node. A doer could choose to follow the strategy proposed by the planner or evaluates his action based on his current expected payoff function. However, when a doer is not sure about his current position, i.e., the decision maker is at a non-singleton information set, the doer uses the planner’s strategy to deduce his current position. The reason why we introduce a planner into our multiself approach is to answer the question, what the decision maker at the ex-ante stage would like to do if he predicts to deviate during execution. The approach includes two kinds of decision makers, one is identical to the one in conventional decision theory, and the other one is the decision maker who takes future deviations into consideration when he chooses a strategy before the play.

Piccione & Rubinstein (1997b) mention that a game should be described as a presentation of the way it is perceived by the players. From such a viewpoint, we could interpret our approach into a manager-agents model. Our model applies to the situation where a team composes of a manager and various agents needs to complete a task together. As a team, the manager and the agents have common payoffs assigned to each outcome. Thus, they share identical preference over outcomes. The planner in our model represents the manager and the doers represent the agents. Each agent takes charge of one section of the whole task. During the process of task completion, some work might not have an immediate effect which makes others difficult to know whether the work has been done or which action was taken. Once the transfer of information among the agents cannot be guaranteed, the agents would be vague about the possible outcomes generated by their current behaviours. Correspondingly, in the decision problem tree, the doers which represent the confused agents might be located in the same non-singleton information set. For example, a task needs two final examinations before it is handed in and different results of examinations have different outcomes. As shown in Figure 3.1, node e_1 represents the first examiner, and node e_2 is the second one. The action E represents that the file fails to pass and C represents it passes. If the first examiner chooses to pass, the task will be examined by the second examiner. If the first examiner chooses to fail, the task will be redone totally. If the second examiner chooses to pass, the task is finished. If the second examiner chooses to fail, the task will be examined by a third examiner (which causes extra time costs even if the task is passed by the third examiner). The standard that whether the task is passed is subjective instead of objective. Additionally, the result of the second examiner relies on the first examiner's result. Then, once the examiner loses the information about his order to examine, the pass rate of the tasks might be different from the prediction before its starts. Correspondingly, the nodes in the decision problem tree which represent the two examiners are located in the same information set (information set I in Figure 3.1).

Not only the example we describe above, but also the situation where each agent separately completes a part of repeated work and is lack conversations with each other, is hard to model by a standard rational agent model. Firstly, an agent in a standard rational model knows the strategy of other agents, even if he does not know other agents' actions. However, like in the example above, every examiner is not able to know the other one's strategy due to the nature of absentmindedness. Once one of the examiners knows the strategy of the other one, or even if he only knows the fact that the other examiner has acted or not. This examiner would know he is the first one or the second one. Besides that, the strategy of an agent in absentmindedness determines not only his behavioural rule but also his position belief in a non-singleton information set. If the strategy of the first examiner is σ to pass, according to consistent belief in conventional game theory, the relative probability of being the first examiner is $1/(1+\sigma)$. It is hard to be modelled in a standard agent model. The issue that an agent's strategy determines his position belief and reversely, the position belief influences his choice of strategy has been solved in our approach by introducing a planner. The doers form their position beliefs based on the planner's strategy. It separates the processes of belief forming and strategy choosing. It is also consistent with the nature of absentmindedness that neither the strategy of the other doers in the same information set with absentmindedness nor the fact whether the other doers have acted can be delivered to the doer.

In the manager-agents model, the manager makes an optimal action plan to complete the task. The action plan should be chosen before the task. The manager's chosen optimal action plan corresponds to the ex-optimal strategy chosen by the planner in the decision problem. The manager clearly knows he cannot make any actual moves. The reason for it could be the manager does not master any technical operation. The outcome of their task is caused by the behaviours of doers instead of the manager. Thus, the manager cares about the agents' actual behaviours. The manager might choose different optimal action plans according to his different beliefs about how agents behave. The manager probably chooses different strategies when he knows the agents might deviate and when he believes the agents strictly follow his strategy.

Since the manager's choice of optimal action plan depends on his beliefs about how agents behave, correspondingly, the ex-ante optimal strategy chosen by the planner in the decision problem also depends on his beliefs about the principle that doers follow when they are called to act. The planner's belief on the doers' behaviours might be various. We analyse two of them in our approach. The planner could believe that the doers will follow whatever strategy he chooses. We say such a planner is **confident**. A confident planner acts the same as the decision maker in conventional decision theory before the decision problem executes. However, as PR point out, the decision maker might later reconsider his strategy during execution and deviate from his ex-ante chosen strategy. Then, should the decision maker take into consideration the possibility of deviation when he chooses a strategy ex-ante? In other words, should the manager take how agents choose their conditional optimal behaviours into consideration when he chooses an optimal action plan? We say the planner who notices doers would evaluate the decision problem themselves and choose their conditional optimal immediate behaviours is

knowledgable, since such a planner knows more information about how doers behave, compared to a confident planner.

For the aspect of doers, we assume that doers evaluate the decision problem based on their knowledge about the decision problem and their current positions, and choose their conditional optimal one-shot behaviour. It is regardless of the planner's belief about doers' behaviours. The assumption of doers corresponds with the intention to examine the consistency of a strategy during execution.

1.4 All about Psychological Games

We have explained that the planner in our multiself approach could be described as confident or knowledgable, and the distinction between these two kinds of planners is caused by their different beliefs about how doers behave once the decision problem is executed. Therefore, we try to introduce the framework of psychological games into our approach. Before explaining why we incorporate psychological games into our approach, we exhibit the previous research on psychological games.

The researches on psychological games are mainly divided into three aspects. The first aspect is the solution concepts developed in the framework of psychological games, especially for extensive psychological games. In an extensive psychological game, a principle similar to sequential rationality in conventional game theory should be defined to examine the conditional optimality of a strategy at an information set or a history in an extensive form game. Furthermore, psychological games models belief-dependent preference. The way that players update their beliefs along their moves should be carefully described. Even if different types of equilibrium are well defined, their existence is still of great importance. The second aspect is to develop a theory of some psychological factors and construct a proper concrete psychological payoff function which can rationalise individuals' behaviours in some games or experiments where the psychological factor intervenes. It is the application of psychological games in the theoretic field. Researchers are not necessary to run an experiment but to explain why an unexpected outcome occurs in terms of those concrete psychological games, such as the fairness function in Rabin (1993). The third aspect is to apply the theory and corresponding psychological payoff functions in the second aspect to different games or experiments and examine the reasonability of those theories and psychological payoff functions by experimental data. Sometimes the second and third aspects could be included in the same paper.

1.4.1 Psychological Game Framework

Psychological games could be used as a tool to solve games related to psychological factors and interactions among individuals. Even for the same two outcomes, players' preferences relation over them might be different if their beliefs about other players' strategies are different. It enlarges the domain space and expects to have a larger set of equilibria. Although there are no obvious psychological factors included in the utility function, people can model individuals' intentions, emotions and attitudes by either including a parameter representing the psychological factors or defining a section of function consisting of players' beliefs and strategy profile to show the effect of psychological factors and including the function into the conventional utility function.

In classic game theory, a player's preference over outcomes can be presented as a function if the preference satisfies some requirements. Once the utility assigned to each outcome is fixed, the payoff function of a player depends only on the players' strategy profile. The only subjective factor in the payoff function is the player's utility of every outcome. If we substitute the utility with the outcome to which the utility is assigned, the resulting function does not change along with players' minds. However, Geanakoplos et al. (1989) (henceforth GPS) propose the concept of psychological games, in which the player's payoff depends on not only players' strategy profile but also depends on players' beliefs about other players' strategies, players' beliefs on other players' beliefs on others' strategies, and so on. It is motivated by an empty threat. An empty threat could be the case that a player commits to take an action at his future move, but when he arrives at the node where the player moves, he has a conditional better choice of action and then takes the better action instead of the action he commits before. However, a player might take the action against the threat when he reaches the relevant node because of some psychological factor. Such behaviour might be seen as an irrational choice of the player. If incorporate the psychological factor into a player's payoff function, the player's action which leads to the empty threat is not an irrational choice with respect to the payoff function including psychological factors. A player's emotional reaction to others' behaviours is normally dependent on his expectation of others' strategies. A classic payoff function cannot perfectly analyse such emotion-relevant scenarios. Thus, psychological games include the beliefs of players as the expectation of a player on other players' strategies. The preferences presented by psychological games somehow belong to other-regarding preferences since they also take others' intensions or attitudes into consideration.

The belief included in the psychological games could be the ones before a game starts and during execution. To be more detailedly, a player's payoff function should be depended on the hierarchy of beliefs. The first order belief of a player is a probability measure over other players' strategies. The second order belief is a probability measure over other players' beliefs and strategies. The third order belief is a probability measure over other players' second, first order beliefs and strategies, and so on. In a psychological Nash equilibrium, the beliefs should coincide with reality, thus, a player's belief of others' strategies should be identical to the strategies that players actually play. In the meanwhile, each player plays one of his best responses to other players' strategy profile and his beliefs. The authors then, prove the existence of a psychological Nash equilibrium in a normal form psychological game if the players' payoff functions are continuous. A subgame perfect psychological equilibrium and a sequential equilibrium exist in an extensive psychological game if players' payoff functions are continuous.

Battigalli & Dufwenberg (2009) (henceforth BD) propose there are several issues which are not covered in extensive psychological games developed in GPS. Only initial beliefs are included in the extensive psychological games in GPS while in BD, they allow updated beliefs in the psychological payoff function since it is crucial for a model over belief-dependent preferences to learn his latest beliefs at any time. BD include not only the player's beliefs but also other players' beliefs in the psychological payoff function. In this way, not only others' strategies, own beliefs on others' strategies and beliefs but also others' beliefs could determine the payoffs of the player. It might provide a better way to model others' emotions and intensions by including their beliefs in their own payoff function. The framework in GPS implies the players' strategies could influence the payoffs by the outcomes to which a strategy profile leads, while BD allows the players' preferences to depend on strategies. In this way, the belief-dependent preferences are extended to strategy and belief-dependent preferences. In BD, not only the existence of equilibria has been discussed but also the situations in which no equilibria exist. They define the psychological sequential equilibrium and prove the existence of such an equilibrium exists as long as the psychological payoff functions are continuous. Besides the major contribution, BD also analyse the work of randomisation in extensive psychological games, extensive psychological games with incomplete information, mutiself approach, and so on.

1.4.2 Fairness and Reciprocity

Fairness is modelled in Rabin (1993). It is a famous application of psychological games. They provide evidence about several interesting facts. People would like to sacrifice their own payoffs to help kind people and punish unkind people. The motivation behind these behaviours becomes more obvious when the amount of payoffs they would like to take out for reward or punishment is relatively smaller. He defines the highest payoffs and lowest payoffs which are Pareto efficient, and thus defines an equitable payoff by assigning equal weight to the highest and lowest payoffs. For player i and j, player i's kindness to player j is evaluated by the difference between the player j's payoff conditional on i's first order belief of j's action and the equitable payoff divided by the difference between the player j's highest payoff and lowest payoff. Thus, player i is kind to player j if the payoff of player j is larger than j's equitable payoff. In the meanwhile, the player i's belief of the kindness of player j is defined as the difference between player i's payoff conditional on his first order belief of what action j takes and his second order belief about the first order belief of player j about what action he would take and i's equitable payoff based on his second order belief divided by the difference between his highest payoff based on his second order belief and lowest payoff based on his second order belief. Then, player i believes j is kind to him if his payoff is higher than his equitable payoff. Then, the player's psychological payoff function composes his material payoff and the payoff brings by his idea of fairness and kindness. The product of a player' i's belief of the kindness of player j and the player i's kindness to player j appears in the psychological part of the payoff. Thus, the player i's payoff increases by being kind to j if he believes j is kind to him, and reversely, being unkind to j if he believes j is not kind to him.

Rabin (1993) also defines the notion of fairness equilibrium on his psychological payoff function that both players do their best responses and the first and the second order of belief is consistent with the real strategy profile. They also define the concepts of mutual-max and mutual-min. A strategy profile is mutual-max if both players choose the action maximising the other player's payoff and mutual-min if both players choose the action minimising the other player's payoff. Thus, a strategy profile is a fairness equilibrium if it is either mutual-max or mutual-min, and if it is a Nash equilibrium. They also demonstrate the existence of a weakly negative fairness equilibrium in which a strategy profile is weakly negative if the kindness functions for both players are smaller than or equal to 0. Other relevant propositions are also proposed.

Falk & Fischbacher (2006) model reciprocity by psychological games. Reciprocity indicates that individuals would like to reward good behaviours and punish bad behaviours. Similar to Rabin (1993), the theory developed in the paper provides a way to evaluate people's kindness and include it into a psychological expected payoff function. They run a large number of experiments with respect to many kinds of games and model kindness by the comparison of different behaviours between the scenario where a participant faces a real person and the scenario where a player plays the game with a machine. The question which plays a key role in their paper is how people evaluate whether a past experience implies kindness or unkindness. They explain why inequity aversion is different from reciprocity. The inequity aversion emphasises outcomes, while reciprocity put more weight on the intention behind people's behaviours rather than physical outcomes. Inequity aversion encourages people's behaviours which reduces inequity while punishing the individuals who increase the inequity. Falk & Fischbacher (2006) define a payoff function consisting of material payoffs and reciprocity payoffs. In the reciprocity part, the product of the kindness term and reciprocity term describes that if the player thinks another player's action is kind, he can increase his payoffs by increasing the other player/s payoffs, and if he receives unkindness from the other player's behaviour, his payoffs would increase if he reduces the other player's payoffs. Then, they take the ultimatum game as an example to show the negative reciprocity and the gift-exchange game to show the positive reciprocity. Then, they prove the existence of a reciprocity equilibrium in those two examples.

Besides the paper we describe above, Dufwenberg & Kirchsteiger (2004) extend the theory of reciprocity to extensive games. Different from a normal form game, when a subgame in an extensive game is reached without expectation, the player's belief of the strategy being played might be changed since he will never reach the current position in the game tree if following the initial strategy in his belief. It is similar to the decision maker's way of consideration in the definition of constrained time consistency in Battigalli (1997). The evaluation of kindness might also be influenced by the change of beliefs. It is necessary to define updated beliefs in extensive psychological games. A player is assumed to keep tracking the change in his belief when he reaches each subgame, and thus he will know his latest belief when making a decision according to it. They use a similar kindness function and utility function of players to those defined in Rabin (1993). They define a sequential reciprocity equilibrium in which players do their best responses and their beliefs identical to the actual strategy profile being played. They prove the existence of a sequential reciprocity equilibrium in every psychological game with the utility functions including the kindness functions.

1.4.3 Guilt Aversion

Battigalli & Dufwenberg (2007) explore the application of psychological games in guilt aversion. A player feels guilty if his actions cause a co-player to receive less than he expected. They define two concepts of guilt aversion. The first one is simple guilt, presented as a payoff function which is the difference between the player's material payoff minus the product of a parameter which reflects the player's guilt sensitivity and a function which indicates his behaviour's effect on how much he lets other players down. In the second version of guilt aversion, the player instead of feeling guilty when he lets others down, cares more about how much others think he intends to let them down or how much others blame him. They apply the two definitions of guilt aversion to a two-player game in which players move simultaneously. If the chance player does not participate in the game, the sequential equilibrium of the psychological games with the two definitions of guilt aversion coincides with each other.

Attanasi & Nagel (2008) apply the two main theoretical theories, guilt aversion, and reciprocity into an example of a trust game to explain the general theoretical framework of psychological games. They also answer some subtle ambiguities regarding the features of the psychological games framework, such as how to define a psychological game is in the setting of complete information since many unknown parameters involved in the psychological payoff function, how to differentiate psychological games and classic games with incomplete information. The answer to the latter question is that the payoffs of a classic game with incomplete information depend on players' strategies and exogenous parameters representing the state of the world. Those factors are invariant with players' subjective perspectives and players are asymmetrically informed of those factors. On the contrary, in psychological games, parameters relevant to players' beliefs in the payoff function are endogenous variables. Attanasi & Nagel (2008) compute the psychological sequential equilibria which are developed by

Battigalli & Dufwenberg (2007) in the trust game with the theory of reciprocity and guilt aversion and provide a way to explore psychological games with incomplete information.

Recent experimental work by Khalmetski et al. (2015) and Ederer & Stremitzer (2017) also provides evidences which is consistent with guilt aversion. Put together these results and it provides evidence that belief-dependent guilt allows cooperation in a manner consistent with the predictions of guilt aversion theory (Kawagoe & Narita 2014). Kawagoe and Narita try to answer the question: "Is it the case that the more others believe that one believes that they take trustworthy actions, the more often they will adopt trustworthy behaviour?" The hypothesis of guilt aversion indicates that people will feel guilty if they betray others' expectation. Naturally, the higher others' beliefs on an individual's belief are, the more in the possibility that one's trust could be rewarded by others' trustworthiness.

1.4.4 Brief Literature on Other Issues

Morrison & Rutström (2002) explore whether it is possible that psychological games are relevant to the resulting comparison between individuals' prior beliefs about others' behaviours and their actual behaviours which are revealed once games start. It is doubtful whether the emotions, such as surprise or disappointment, that appear when individuals confirm the consequences of players' strategy interactions, would have effects on their payoffs. They design and implement an experiment that allows for belief elicitation in an investment game, where participants are totally self-interested and calm. The assumption indicates there might be little trust or reciprocity in the game. They elicit the truster's first order beliefs before he makes his decision and the second order beliefs before he knows the decisions made by his partner. The results of their experiments support the theory of reciprocity in Rabin (1993) since the trustees in the experiments repay trusters as the feedback of their kindness which indicates the existence of reciprocal behaviours. Their results of the experiments also are consistent with the inequality aversion model in Fehr & Schmidt (1999). However, there is a strong correlation between beliefs and amounts of money which is sent back to trusters. It seems not to be ably explained by the model of inequality aversion.

Dufwenberg & Gneezy (2000) analyse the relevance of trust responsiveness in a trust game. In the game, a truster may take an amount of money between 0 and 20 units, otherwise, they could also leave them and let a trustee takes away an amount of money between 0 and 20. The truster and the trustee simultaneously make their choices, Then, elicit the first-order beliefs from the truster and the second-order beliefs from the trustee. They conclude that trust responsiveness should hold even if belief transmission did not take place. Trust responsiveness predicts that trustees who have higher guesses will be more likely to realise trust.

So far, we could see that even though the game is called psychological games, there are few theories to model emotions through psychological games. Battigalli et al. (2019) develop a general model of frustration and anger in games. Players are assumed to be frustrated if they meet an outcome which is worse than their expectation, then players will put less weight on their co-players material payoffs and increase their psychological payoff from the aggressive behaviour. Define frustration the difference between the best outcome a player can still receive in the game and the material payoff that the player had initially expected. They associate anger with frustration, modelling anger as a function of the blame that one player places on another player for his frustration. If there are two players i and j in the game. Player i receive frustration from the difference between the highest payoff that he still can get and the payoffs he expects to receive at the beginning of the game. Player i blames player j because of his frustration and player j feels angry because of being blamed. Thus, the player j's preference over the outcomes changes as the value of the factor which indicates his anger changes. If the player j feels angry, he prefers an action which leads to an outcome that punishes player i to the action which is optimal when he was not angry. Once the utility functions of both players are defined, they could apply them to multiple types of games and design experiments to verify whether the psychological payoff functions they develop are consistent with the experimental results.

To sum up, the research on the application of psychological games could be developed in many fields, especially those in which emotions and interaction are involved. Besides the papers we mention above, there are numbers of papers which discuss relevant issues. Battigalli & Dufwenberg (2020) could be a good paper to learn the brief results of research on psychological games. However, there are few works present advanced techniques to solve the issues regarding the basic framework of psychological games. In comparison with the theoretical improvements, many researchers design and run experiments to explore the effect of some psychological factors on people's behaviours and propose their theories to describe and rationalise "irrational" behaviours.

1.4.5 Why We Incorporate Psychological Games into Our Approach?

As we stated above, psychological games are used to describe belief-dependent preferences. Belief-dependent preference presents that individual's preferences over objects are different if their beliefs are different. Here, the belief refers to an individual's belief about others' behaviours and their beliefs. Individuals' emotions and other psychological factors which influence their behaviour reflect in their different optimal behaviours when they hold different beliefs about others' behaviours. However, it does not mean psychological games could only explain psychological factors relevant issues. The essence of psychological games is to describe the scenario that people might act differently from the way which is described in conventional game theory even if we always assume people are rational when they choose what to do. Individual's different optimal behaviours might result in new equilibria which differ from those in conventional game theory appears. The introduction of belief-dependent preference rationalises the occurrence of those "irrational" behaviours in conventional game theory. Thus, under the assumption of rationality, as long as a behaviour or an equilibrium could be explained by the different beliefs of players, we could explain it by the framework of psychological games no matter whether it is relevant to emotion or other psychological factors.

We choose to use multiself approach to explain imperfect recall (absentmindedness). In our analysis, a decision maker is modelled as the planner self and the doer selves. When the planner chooses a strategy for the decision problem, he needs to consider how doers will behave later in the stage of execution since he is not the one whose actions could cause real outcomes. The planner's "knowledge" (or belief) about how the doers deal with his ex-ante strategy influence much on his choice of it. Thus, a new equilibrium between the planner and doers could exist compared to those in multiself analysis of conventional decision problem theory. The new equilibrium could exist in the form of different strategies between the planner and the collective of doers. If we speak of a strategy which satisfies time consistency in conventional decision theory, a different time consistent strategy might exist due to the planner's different beliefs on doers' actions.

In fact, instead of people's emotions and interactions, psychological games in our model are used to describe different degrees of knowledge about people themselves. The appearance of imperfect recall (absentmindedness) makes a decision maker might act differently at the stage of execution from the strategy he chooses at the stage of planning. The decision maker might take some actions or change his way of evaluating the decision problem at the stage of planning if he notices possible deviation appearance at execution stage. According to his knowledge about himself, the decision maker at the stage of planning could form different beliefs about how he behaves during execution. Then, according to his different beliefs about his behaviours during the play, he might choose a different strategy ex-ante. A confident decision maker who believes that he would strictly follow whatever strategy he has chosen ex-ante when the decision problem is executed. When the decision maker chooses a strategy ex-ante, he could choose the one which is most beneficial (the one maximising the ex-ante expected payoffs) for the whole play. The confident decision maker believes his ex-ante chosen strategy is a credible commitment. Once, the equilibrium between the planner and the doers exists, the corresponding equilibrium strategy is time consistent in the standard of conventional decision theory. Compared, the knowledgable decision maker knows he will reevaluate the decision problem during the stage of execution, and the reevaluation could cause his deviation from the ex-ante chosen strategy. The decision maker forms belief on how he will take advantage of his ex-ante strategy. Being a rational decision maker, he should choose a strategy which helps him in the execution stage most. In our model, we assume the knowledgable decision maker knows how he will behave in the stage of execution. Whenever the decision maker meets imperfect recall and cannot confirm his current position in the decision problem tree, the decision maker uses the ex-ante chosen strategy as evidence to deduce his precise position. However, we never claim the decision maker must form beliefs the same as the knowledgable decision maker, it is only one possibility from thousands of beliefs. For example, the decision maker is not necessary to reconsider the decision problem at every node. Another belief could be the decision maker has some possibility to deviate although he reevaluates at every node along the play.

The two types of planners, confident and knowledgable, represent the different beliefs of the decision maker in ex-ante stage about how he will behave during execution. Thus, the analysis of both types of planners composes the whole structure of a psychological game. It is easy to see that the psychological game used in our approach is not a typical one discussed in most psychological games literature. Firstly, we cannot guarantee the planner's different beliefs could be written in a unique function form. We could know from the psychological game relevant analysis of regret, reciprocity and fairness that, in their framework, only the value of some parameters changes if the decision maker's belief changes. However, the definition of psychological games in GPS does not require the psychological payoff function must be the identical functional form. Except that, in most of the research on guilt, regret and other psychological factors, the psychological payoff function includes the material payoff part and psychological payoff part. However, in our approach, we use the expected payoff function as the psychological payoff function of the planner and doers. There is no distinct part between material payoff and psychological payoff. It is because our approach focuses more on strategy-change issues, i.e., whether the decision maker changes his ex-ante chosen strategy when he reconsiders during the decision problem processes. The third aspect that our approach is different from the framework of conventional psychological games is that our approach does not describe individuals' emotions but the knowledge about themselves, and how they deal with behaving. Since our approach is to adopt a new technique to analyse the inconsistency between ex-ante optimality and execution conditional optimality, the decision maker in the stage of execution must choose his immediate behaviour which maximises the conditional expected payoffs but not just follows the ex-ante chosen strategy. However, if we apply our approach to the manager-agents model, the decision maker's knowledge about himself becomes the manager's knowledge about the agents. We extend the manager-agents model in the way that the agents could choose their behaviours according to different ways when dealing with conflict situations (follow the ex-ante strategy or deviate to the conditionally optimal action). Then, how the agents deal with conflict situations becomes the agents' internal information but not the common information. If the way that agents behave becomes common knowledge between agents and the manager and even if there is prior about it, the problem transfers into a bayesian game.

1.5 Revisit Absentminded Driver Paradox

We use absentminded driver paradox as an example to illustrate our approach and compare the equilibria results of different beliefs of the planner.

1.5.1 Standard Absentminded Driver Paradox

We need to define notation first,

a) The strategy and utility function for the driver is $s_r = (r, r)$ and u.

b) The strategy of the planner is $s_t = (t, t)$. There is one information set in this example. The planner suppose both the doers at e_1 and e_2 know the planner's strategy s_t.The planner's utility function is U

c) The actions of the doers at intersection e_1 and e_2 are p_1, p_2 respectively. The utility function for them is V_1 and V_2 respectively.

d) The ith order belief of the planner is $b^i_{\cdot,pl}$. Here is an implicant that the related beliefs are identical to all doers. All the variables defined refer to the probability of the choice to continue.

Doers

For the driver at one intersection in formation set I, his conditional expected payoff function is

$$Eu(t|I;s_r,t) = \frac{1}{t+1}(4t(1-t)+t^2) + \frac{t}{t+1}(4(1-t)+t),$$

given his strategy s_r. Now we explain the conditional expected payoff function. The s_r in $Eu(t|I;s_r,t)$ represents the driver at an intersection at information set I forms his belief over probabilities of being at each node in I by a belief system consistent with the strategy s_r. The first t in $Eu(t|I;s_r,t)$ (the one before the vertical line) describes the decision maker's current one-shot behaviour. It also implies the driver is able to change his current one-shot action. The second t in $Eu(t|I;s_r,t)$ (the one after the vertical line) stands for his behaviour at the other intersection is t.

A doer's strategy is a one-shot behaviour of the driver at the node where the doer is located. The doer at e_1 chooses his strategy p_1 conditional on his belief formed based on the planner's strategy, $s_t = (t,t)$, and the strategy of doer at e_2, p_2. The doer at e_2 chooses his strategy p_2 conditional on his belief formed based on the planner's strategy, $s_t = (t,t)$, and the strategy of doer at e_1, p_1. Then, the payoff functions for doers at e_1 and e_2 are

$$V_1(p_1|s_t,p_2) = Eu(p_1|I;s_t,p_2) = \frac{1}{t+1}(4p_1(1-p_2)+p_1p_2)+\frac{t}{t+1}(4(1-p_1)+p_1)$$

$$V_2(p_2|s_t,p_1) = Eu(p_2|I;s_t,p_1) = \frac{1}{t+1}(4p_2(1-p_1)+p_2p_1)+\frac{t}{t+1}(4(1-p_2)+p_2).$$

We can know that the equilibrium strategy should be symmetric, (a,b) should also be an equilibrium if (b,a) is an equilibrium.

Confident Planner

If the planner is **confident**, he believes that the doers will follow whatever strategy he chooses before the decision problem runs. The confident planner will choose an optimal strategy to reach the highest expected payoff at the planning stage. It is also the ex-ante expected payoff function in conventional decision problem. The planner's payoff function is

$$U(s_t) = Eu(s_t) = t^2+4t(1-t) = 4t-3t^2 = -3(t-\frac{2}{3})^2+\frac{4}{3}.$$

Therefore, the confident planner's optimal strategy is $s_t^\star = (\frac{2}{3},\frac{2}{3})$.

Substitute $s_t^\star = (\frac{2}{3},\frac{2}{3})$ into the doers' functions, it results that

$$V_1(p_1|(\frac{2}{3},\frac{2}{3}),p_2) = \frac{3}{5}(4p_1(1-p_2)+p_1p_2)+\frac{2}{5}(4(1-p_1)+p_1) = \frac{1}{5}\left[(6-9p_2)p_1+8\right],$$
$$V_2(p_2|(\frac{2}{3},\frac{2}{3}),p_1) = \frac{3}{5}(4p_2(1-p_1)+p_2p_1)+\frac{2}{5}(4(1-p_2)+p_2) = \frac{1}{5}\left[(6-9p_1)p_2+8\right],$$

In equilibrium,

$$p_1^\star = \underset{p_1}{argmax} V_1(p_1|(\frac{2}{3},\frac{2}{3}),p_2^\star), \quad p_2^\star = \underset{p_2}{argmax} V_2(p_2|(\frac{2}{3},\frac{2}{3}),p_1^\star).$$

Solve them, we get $(p_1^\star, p_2^\star) = (1,0),(\frac{2}{3},\frac{2}{3})$ or $(0,1)$. Notice that, the doers at e_1 and e_2 are in the identical situation and need to deal with the same optimisation problem. It indicates the doer at e_2 should also assume the behaviour of the one at e_1 is a if the doer at e_2 believes the doer at e_1's strategy is a, vice versa. Then, the equilibria $(0,1)$ and $(1,0)$ are not proper results since it implies the doers have different knowledge on each other's strategy. Therefore, it implies the condition $p_1^\star = p_2^\star$ is naturally required in equilibrium.

The psychological multiself equilibrium with a confident planner $(s_t^\star,(p_1^\star,p_2^\star)) = ((\frac{2}{3},\frac{2}{3}),(\frac{2}{3},\frac{2}{3}))$. It shows that although the doers reevaluates the decision problem at their current position, they seems to follow the confident planner's strategy in the viewpoint of the planner. Since how the doers behave is the internal information of doers themselves, the planner can only observe the doers' behaviours and translate their behaviours with the reference of his own belief on doers.

Knowledgable Planner

For a **knowledgable** planner, he knows doers regard his strategy as the reference for deducing the probability distribution of his current position in an information set. The planner knows the doers are rational and forms his first order belief on doers' strategies that they should do their best responses according to the calculated beliefs by the planner's strategy.

The doers' payoff functions are

$$V_1(p_1|s_t,p_2) = Eu(p_1|I;s_t,p_2) = \frac{1}{t+1}(4p_1(1-p_2)+p_1p_2) + \frac{t}{t+1}(4(1-p_1)+p_1)$$

$$V_2(p_2|s_t,p_1) = Eu(p_2|I;s_t,p_1) = \frac{1}{t+1}(4p_2(1-p_1)+p_2p_1) + \frac{t}{t+1}(4(1-p_2)+p_2).$$

Then, their best responses are

$$p_1^{BR}(s_t) = \underset{p_1}{argmax} V_1(p_1|s_t, p_2), \quad p_2^{BR}(s_t) = \underset{p_2}{argmax} V_2(p_2|s_t, p_1).$$

Solve it, we get $p_1^{BR}(s_t) = p_2^{BR}(s_t) = \frac{4-3t}{3}$ or $p_1^{BR}(s_t) = p_2^{BR}(s_t) = 1$ if $t \in [0, \frac{1}{3}]$. Thus, the planner's beliefs on doer at e_1 and e_2 are respectively $b^1_{p_1,pl}(s_t) = p_1^{BR}(s_t) = \frac{4-3t}{3}$, $b^1_{p_2,pl}(s_t) = p_2^{BR}(s_t) = \frac{4-3t}{3}$ or $b^1_{p_1,pl}(s_t) = p_1^{BR}(s_t) = 1, b^1_{p_2,pl}(s_t) = p_2^{BR}(s_t) = 1$ if $t \in [0, \frac{1}{3}]$. Then substitute the equation to the payoff function of the knowledgable planner,

$$\begin{aligned} U^1(s_t|b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) &= Eu(t|I; b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) \\ &= \frac{1}{t+1}\left(4b^1_{p_1,pl} - 3b^1_{p_1,pl}b^1_{p_2,pl}\right) + \frac{t}{t+1}\left(4 - 3b^1_{p_2,pl}\right) \\ &= \frac{1}{t+1}\left[4(\frac{4-3t}{3}) - 3(\frac{4-3t}{3})^2\right] + \frac{t}{t+1}\left[4 - 3(\frac{4-3t}{3})\right]; \end{aligned} \tag{1.1}$$

or

$$U^1(s_t|1,1) = Eu(s_t|I;1,1) = \frac{1}{t+1}[4 \cdot 1 - 3 \cdot 1^2] + \frac{t}{t+1}[4 - 3 \cdot 1].$$

Then, in equilibrium, $p_1^\star = b^{1,\star}_{p_1,pl} = b^1_{p_1,pl}(s_t^\star)$ and $p_2^\star = b^{1,\star}_{p_2,pl} = b^1_{p_2,pl}(s_t^\star)$, we get $\left((b^{1,\star}_{p_1,pl}, b^{1,\star}_{p_2,pl}); s_t^\star, (p_1^\star, p_2^\star)\right) = \left((\frac{1}{3}, \frac{1}{3}); (1,1), (\frac{1}{3}, \frac{1}{3})\right)$ or $\left((1,1), ([0, \frac{1}{3}], [0, \frac{1}{3}]), (1,1)\right)$.

1.5.2 $x - y - z$ General Version of Absentminded Driver Paradox

Now we extend our original absentminded driver paradox to the version with general $x - y - z$ payoffs attaching to each terminal histories. The decision problem shows as in Figure 3.2.

It is obvious that if x is the largest among three payoffs, the decision maker's optimal strategy is EXIT at both intersections, ending up with the payoff of x, and he will not change his strategy at the information set I. Similarly, if z is the largest among three numbers, the decision maker will stick to CONTINUE at any stage of the decision problem. Therefore, our discussion will exclude these two situations and focus on what will happen if $y > x > z$ and $y > z > x$. Without generality, denote $y' = y - z$, $x' = x - z$ and $z = 0$ when $y > x > z$, then $y' > 0$ and $x' > 0$; denote $y" = y - x$, $z" = z - x$ and $x = 0$ when $y > z > x$, then $y" > 0$ and $z" > 0$.

Situation 1: $y > x > z$

For the doers, the payoff function are

$$V_1(p_1|s_t, p_2) = Eu(p_1|I; s_t, p_2) = \frac{1}{t+1}((1-p_1)x' + p_1(1-p_2)y') + \frac{t}{t+1}(y'(1-p_1)),$$

$$V_2(p_2|s_t, p_1) = Eu(p_2|I; s_t, p_1) = \frac{1}{t+1}((1-p_2)x' + p_2(1-p_1)y') + \frac{t}{t+1}(y'(1-p_2)).$$

For a confident planner, the expected payoff function is

$$U(s_t) = Eu(s_t) = x'(1-t) + y'(1-t)t = -y'(t - \frac{y'-x'}{2y'})^2 + \frac{(y'+x')^2}{4y'}.$$

The payoff function of the confident planner reaches the highest expected payoff $\frac{(y'+x')^2}{4y'}$ as long as $t = \frac{y'-x'}{2y'}$. Substitute it into v_1 and v_2, we get $(s_t^\star, (p_1^\star, p_2^\star)) =$ $((\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}), (\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}))$,$((\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}), (1,0))$ or $((\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}), (0,1))$. Same as analysis in 4.1.1, $p_1^\star = p_2^\star$ is naturally required in equilibrium. Then, the only equilibrium is $(s_t^\star, (p_1^\star, p_2^\star)) = ((\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}), (\frac{y'-x'}{2y'}, \frac{y'-x'}{2y'}))$. Notice that $s_t^\star = (p_1^\star, p_2^\star)$.

The doers' best responses for a planner with the strategy s_t are $p_1^{BR}(s_t) = p_2^{BR}(s_t) = \frac{y'-x'}{y'} - t$ or $p_1^{BR}(s_t) = p_2^{BR}(s_t) = 0$ if $t \in [\frac{y'-x'}{y'}, 1]$. Then, the first order belief of the planner on doers' strategies are $b^1_{p_1,pl}(s_t) = b^1_{p_2,pl}(s_t) = \frac{y'-x'}{y'} - t$ or $b^1_{p_1,pl}(s_t) = b^1_{p_2,pl}(s_t) = 0$ if $t \in [\frac{y'-x'}{y'}, 1]$. The payoff function of the knowledgable planner is

$$
\begin{aligned}
U^1(s_t|b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) &= Eu(t|I; b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) \\
&= \frac{1}{t+1}[(1-b^1_{p_1,pl})x' + b^1_{p_1,pl}(1-b^1_{p_2,pl})y'] + \frac{t}{t+1}[y'(1-b^1_{p_2,pl})] \\
&= \frac{1}{t+1}[(2y'-x')(-t+\frac{y'-x'}{y'}) - y'(-t+\frac{y'-x'}{y'})^2 + x' - y'] \\
&+ y'(1-(-t+\frac{y'-x'}{y'}))
\end{aligned}
$$

or

$$
U^1(s_t|0,0) = Eu^1(t|I;0,0) = \frac{1}{t+1}[(2y'-x')\cdot 0 - y'\cdot 0^2 + x' - y'] + y'(1-0) = y' - \frac{y'-x'}{t+1}.
$$

Solve them, we conclude the equilibria are $\left((b^{1,\star}_{p_1,pl}, b^{1,\star}_{p_2,pl}); s^\star_t, (p^\star_1, p^\star_2)\right) = ((\frac{y'-x'}{y'}, \frac{y'-x'}{y'}); (0,0), (\frac{y'-x'}{y'}, \frac{y'-x'}{y'}))$ or $((0,0);(1,1),(0,0))$.

Situation 2: $y > z > x$

For the doers, the payoff function are

$$
\begin{aligned}
V_1(p_1|s_t, p_2) = Eu(p_1|I; s_t, p_2) &= \frac{1}{t+1}p_1(y''(1-p_2) + z''p_2) + \frac{t}{t+1}[y''(1-p_1) + z''p_1], \\
V_2(p_2|s_t, p_1) = Eu(p_2|I; s_t, p_1) &= \frac{1}{t+1}p_2(y''(1-p_1) + z''p_1) + \frac{t}{t+1}[y''(1-p_2) + z''p_2],
\end{aligned}
$$

For a confident planner, the payoff function is

$$
U(s_t) = Eu(s_t) = t(z''t + y''(1-t)) = (z''-y'')(t - \frac{y''}{2(y''-z'')})^2 + \frac{y''^2}{4(y''-z'')}.
$$

The payoff function reaches the highest expected payoff $\frac{y"^2}{4(y"-z")}$ as long as $t = \frac{y"}{2(y"-z")}$. Substitute it into v_1 and v_2 and $p_1^\star = p_2^\star$ in equilibrium, then the only equilibrium is $(s_t^\star, (p_1^\star, p_2^\star)) = \Big((\frac{y"}{2(y"-z")}, \frac{y"}{2(y"-z")}), (\frac{y"}{2(y"-z")}, \frac{y"}{2(y"-z")}) \Big)$. Notice that $t^\star = p_1^\star = p_2^\star$.

The best responses of doers to a planner with strategy s_t are $p_1^{BR}(s_t) = p_2^{BR}(s_t) = 1$ if $t < \frac{z"}{y"-z"}$, $p_1^{BR}(s_t) = p_2^{BR}(s_t) = -t + \frac{y"}{y"-z"}$, or $p_1^{BR}(s_t) = p_2^{BR}(s_t) = 0$ if $t > \frac{z"}{y"-z"}$. Notice that $\frac{z"}{y"-z"} > 1$ if $z" > \frac{1}{2}y"$, the equilibrium $p_1^\star(t) = p_2^\star(t) = 0$ does not exist. Then, the payoff function of the knowledgable planner is

$$\begin{aligned}
U^1(s_t|b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) &= Eu(t|I; b^1_{p_1,pl}(s_t), b^1_{p_2,pl}(s_t)) \\
&= \frac{1}{t+1}[(z"-y")(-t+\frac{y"}{y"-z"})^2 + (2y"-z")(-t+\frac{y"}{y"-z"}) - y"] \\
&+ [y" + (z"-y")(-t+\frac{y"}{y"-z"})]
\end{aligned}$$

or

$$U^1(s_t|1,1) = Eu(t|I;1,1) = \frac{1}{t+1}[(z"-y")\cdot 1^2 + (2y"-z")\cdot 1 - y"] + [y" + (z"-y")\cdot 1]$$

or

$$U^1(s_t|0,0) = Eu(t|I;0,0) = \frac{1}{t+1}[(z"-y")\cdot 0^2 + (2y"-z")\cdot 0 - y"] + [y" + (z"-y")\cdot 0]$$

Solve them, we conclude the equilibria are $\Big((b^{1,\star}_{p_1,pl}, b^{1,\star}_{p_2,pl}); s_t^\star, (p_1^\star, p_2^\star) \Big) = ((1,1);(K,K),(1,1))$ or $((\frac{z"}{y"-z"}, \frac{z"}{y"-z"});(1,1),(\frac{z"}{y"-z"}, \frac{z"}{y"-z"}))$. Here, $K := [0, \frac{z"}{y"-z"}] \cap [0,1]$.

1.5.3 Recap Equilibrium Results

Our model separates the analysis of a decision problem with imperfect recall into two steps, planing and execution. The decision maker at ex-ante stage is regarded as a self, a planner. The planner proposes a strategy based on his beliefs of how the decision maker behaves at execution stage. The decision maker at execution stage is divided into several selves, one self at one history, called doers. Assume that a decision maker with imperfect recall is not able to remember his updated strategy even if he deviates to a new strategy before arriving at the current history. In that case, the decision maker is allowed to change his behavioural rule restricted to the current history. He will not deviate to a full strategy which is beneficial for the rest of the decision problem since he knows the future he could not be informed of his new decisions at the current information set. The decision maker only remembers the ex-ante chosen strategy, i.e., the strategy of the planner, and views the strategy as evidence to form his belief on his position in the decision problem. Although we cannot see any notation of beliefs in the calculation of the psychological multiself approach with a confident planner, there is an implicit assumption that the planner believes doers will follow his strategy. We actually have substituted $t = b^1_{p_1,pl} = b^1_{p_2,pl}$, which should be the equilibrium requirements, into the planner's utility function for convenience. Thus, the planner's utility function is a psychological utility function although we do not write it in such a format.

From the results, we find that the collective equilibrium strategy of doers is the same as the equilibrium strategy of a confident planner and they are identical to the optimal strategy of the absentminded driver example. It is reasonable to speculate that, in a general decision problem with imperfect recall, whether the optimal strategy is the same between the equilibrium strategy of the planner and the collective equilibrium strategy of doers. The results of the psychological multiself approach with a knowledgable planner show that the collective equilibrium strategy of doers is different from the equilibrium strategy of the knowledgable planner. If the fact is also tenable in a general case, it verifies a rule that people will receive what they believe. The confident planner believes doers must follow his strategy and no doer deviates in equilibrium while the knowledgable planner believes doers will deviate during execution, which results in different equilibria between the planner and doers.

We try to use a description which is similar to the two-identical-agent formulation in Gilboa (1997) to explain why $p_1^\star = p_2^\star$ in equilibrium. As it is shown in Figure 3.3,

the branches from history O present the decision maker at e_1 and e_2 respectively and β is the probability that the decision maker is at intersection e_1. The two branches represent their current actual position in the information set which can be observed by an outside party for certain. We call the decision maker at e_1 doer 1 and the decision maker at e_2 doer 2. The two doers meet the identical situation. Denote p_{e_i} the doer i's current move and q_{e_i} the behaviour if he reaches the same undistinguishable intersection again. Then, $p^\star_{e_1} = p^\star_{e_2}$ since the decision maker deals with the same optimality problem at both intersections. Doer 1 cannot differentiate which intersection he reaches, i.e., he cannot distinguish history A or history N. His current move p_{e_1} is followed by uncontrolled move q_{e_1} if he is at A. Similarly, the doer 2 does not know he is at history B or history M. If he is at history M, his current move is p_{e_2} which should be the same as q_{e_1} since M is the history that doer 1 reaches after executing his strategy p_{e_1}. Thus, $p^\star_{e_2} = q^\star_{e_1}$ and similarly, $p^\star_{e_1} = q^\star_{e_2}$. In conclusion, we have $q^\star_{e_2} = p^\star_{e_1} = p^\star_{e_2} = q^\star_{e_1}$, i.e., $p_1 = p_2$ in our model.

Both doers' beliefs on the information set are the same, i.e. $p(A) = p(B) = \frac{1}{t+1}$ and $p(M) = p(N) = \frac{t}{t+1}$ if the planner's strategy is $s_t = (t,t)$. The probability distribution depends on neither themselves' nor the other doer' behaviours, which successfully makes the independence of beliefs and actions. Besides, the probability distribution is not influenced by β, the probability of the decision maker's actual position in the information set I, although most works regard $\beta = 0.5$ as a natural assumption, as the analysis in Gilboa (1997). It is because the decision maker is assumed to not be able to distinguish any two situations which are described as the nodes in the same information sets by observing the surrounding environment. If an individual does not form prior probability distribution among different situations, he would assign equal probability to each situation.

Not only in this example but also in any other decision problem that includes non-singleton information sets, behaviours at different histories in the same information set must be identical in equilibrium. It is because the equilibrium strategy should be identical and symmetric. Therefore, in a multiself approach, it is unreasonable to assume the decision maker deviates at solely one history from a non-singleton information set. The selves are able to at most deviate to a same new behavioural rule in equilibrium. It is different from the assumption that the decision maker remembers the new action plan when the same information set occurs again in the future. In our model, the decision maker evaluates his current action assuming he behaves arbitrarily at other histories in the same information set.

If a strategy p is modified multiself consistent in the absentminded driver example, then the strategy satisfies

$$p = argmax_q \left(\frac{1}{p+1}(4q(1-p)+p) + \frac{p}{p+1}(4(1-q)+q) \right) = argmax_q \left(\frac{(4-6p)q+4p}{p+1} \right)$$

Solve it, the strategy $p = \frac{2}{3}$ is the only modified multiself consistent strategy. It is easy to know the strategy is the same as optimal strategy, collective equilibrium of doers and equilibrium strategy of the planner in the psychological multiself method with confident planner. Besides this equilibrium result, the psychological multiself approach reaches more results if the planner's belief changes. Thus, the psychological multiself concept extend the equilibrium space.

1.6 Psychological Multiself Approach

In this section, we would like to develop our psychological multiself approach formally. First, we collects definitions and notation about some concepts used in the literature on extensive games and psychological games. The next subsection is to carefully define the framework of psychological multiself approach.

1.6.1 Basic Definitions and Notation

A formal definition of an extensive decision problem is established in this part. Then, the definitions of relevant concepts are explained based on given notation.

Decision Problems in Extensive Form

The presentation follows Osborne & Rubinstein (1994). An extensive decision problem $\Gamma = < H, u, C, \rho, X >$ has a structure as follows.

• A set H of action sequences (We only discuss finite decision problems which have finite moves.) which satisfies two properties as follow,

1. The empty sequence $\emptyset$ is an element of H.

2. If $h = (a_i)_{i=1,\dots,K} \in H$, then for any $L \leqslant K$, $(a_i)_{i=1,\dots L} \in H$. We call the later is a **subhistory** of h.

An element in H is called a **history** h, a_i is defined as the i-th action in any history h, and H is the set of histories. A history $h = (a_i)_{i=1,\dots,K} \in H$ is **terminal** if there is no a_{K+1} such that $(a_i)_{i=1,\dots,K+1} \in H$. The set of available actions at any nonterminal history h is $A(h) := \{a : (h,a) \in H\}$. The set of terminal histories is denoted by Z, $Z := \{h \in H | \nexists a \in A, s.t.(h,a) \in H\}$. Denote $h' \preceq h$, if h' is a subhistory of h; $h' \prec h$, if h' is a subhistory of h but different from h. If we present an extensive decision problem in a diagrammatic form, the history set H can be drawn as a tree with the empty sequence $\emptyset$ as the root and the $\prec$-relation as edges. If the set H of possible histories is finite then the decision problem is finite. We only discuss **finite** decision problems in this paper.

• A **utility (payoff) function** $u : Z \to \Re$ which is defined over terminal histories and assigns each of them a utility (or payoff) number. Preferences can be presented as lotteries on terminal histories and satisfy the VNM assumption.

• C is a subset of H. The chance player moves after histories in C.

• A function ρ indicates a probability distribution among actions in $A(h)$ whenever $h \in C$ which represents the decision maker's knowledge about the chance player's behaviour. The function is assumed to be strictly positive

• The set of information sets, denoted by $\mathscr{X}$, is a partition of $D = H - Z - C$. $\mathscr{X}$ represents the decision maker's information partition. The decision maker has the same knowledge of the decision problem and its past play at any history belonging to the same **information set**, thus, the set of available actions attaching to each history in an information set should be identical to each other, namely $A(h) = A(h')$ whenever h and h' are in the same information set $X \in \mathscr{X}$. We can write the set of actions which are available at history in X by $A(X)$.

Imperfect Recall and Absentmindedness

The **experience**, denoted by $exp(h)$, of a decision maker at a history h in D is the sequence of information sets and actions which the decision maker undergoes along the history h. Define the last element in the sequence exp(h) is the information set which contains h. A decision problem has **perfect recall** if for any two histories, $h, h' \in D$, which belong to the same information set, the corresponding experiences are identical, i.e., $exp(h) = exp(h')$. Therefore, in a decision problem with perfect recall, the decision maker remembers all past the information sets he has met and the actions he has taken. A decision problem contains information sets which violate the above condition is said to be a decision problem with imperfect recall. A decision maker is said to have imperfect recall if he forgets his previous actions or previous acquired information.

Decision problems with imperfect recall can be categorised into mainly three situations. We use three examples to explain these aspects. In example 1 (see Figure 3.4), the decision maker forgets his previous move at e_1 and cannot distinguish between e_2 and e_3 when he is going to choose his second move. In example 2 (see Figure 3.5), the decision maker knows the move of chance player at e_1 and e_2 but forgets this information at the information set including e_3 and e_4. In example 3 (see Figure 3.1), the decision maker is not able to differentiate two histories e_1 and e_2 in the same path, which refers to absentmindedness. A decision problem Γ exhibits **absentmindedness** if there are two histories h and h' along a path such that h' is a subhistory of h ($h' \neq h$) and both belong to the same information set, i.e., $\exists$ histories h', h and information set $X \in \mathscr{X}$, s.t. $h' \prec h \in X$. The difference between the first situation and the absentmindedness is that the decision maker remembers he has made a decision but forgets which decision is selected while absentmindedness refers to the fact that the decision maker forgets whether he has made a decision or not.

We use another viewpoint to explain the difference among the three situations. To discuss the issues regarding the execution of a decision problem, especially when the decision maker's strategy involves randomisation among pure behaviours, we need to distinguish the decision maker's knowledge of the strategy and the action which is the execution result of a strategy at each node. For example, in Figure 3.4, the decision maker at e_2 might know the strategy he proposes at e_1 indicates to execute the actions L and R with equal probability. However, he does not know the execution result of this strategy since he is now in the non-singleton information set

K and cannot distinguish the nodes e_2 and e_3.

Now we allow the decision maker to record the strategy he executes in the midst of the decision problem. In the first situation (see Figure 3.4), the decision maker is able to record both the strategies and the nodes where those strategies are executed. However, the decision problem might not know the execution result of those strategies. In the second situation (see Figure 3.5), the decision maker is able to record the strategies that he executes but doer not write down at which node a strategy is implemented. For example, the decision maker at node e_3 might know the strategy he executes at e_1 or e_2. He knows that the execution result of his strategy at node e_1 or e_2 is action D but does not know at which node he implements the strategy. In the situation of absentmindedness, the decision maker is not able to record either the strategies he has implemented or the nodes he has passed. Once either the strategies or the nodes are written down, the decision maker knows his precise location in the decision problem tree.

Strategies

A **strategy**, $\sigma \in \Sigma$, is a function which assigns to every history $h \in D$ an element (**pure strategy**) or a probability distribution (**behavioural strategy**) of $A(h)$. Σ is the set of behavioural strategies. There is an underlying restriction that the decision maker behaves the same at any histories within an information set, i.e., $\sigma(h) = \sigma(h')$ if h and h' are in the same information set. A **mixed strategy** is a lottery over the set of pure strategies. In any decision problem, a mixed strategy can never obtain strictly higher expected utility than any pure strategies. In a behavioural strategy, only one random device (random behavioural rule) is used at an information set, but the device should be independently realised if the information set is visited more than once (which is called "absentmindedness").

At an information set X or history h, the probability assigned by a behavioural strategy σ to the action $a \in A(X)$ or $a \in A(h)$ is denoted by $\sigma(a|X)$, $\sigma(a|h)$. The probability of reaching history h' conditional on h given σ is denoted by $p(h'|h,\sigma)$. Thus $p(h|\emptyset;\sigma)$ is the prior probability of history h given σ. Let

$$p(X|\emptyset;\sigma) := \sum_{h\in X} p(h|\emptyset;\sigma).$$

A strategy σ is ***ex-ante optimal*** if it maximises the expected utility function,

$$Eu(\sigma) = \sum_{z\in Z} p(z|\emptyset;\sigma)u(z).$$

According to Kuhn's theorem, in a game with perfect recall, any mixed strategy has a behavioural strategy which induces the same probability measure over terminal histories, thus obtains the same expected payoffs, and vice versa. It is not true for a game with imperfect recall. We still use Absentminded Driver paradox, a one-player game, as an example to prove it. As we analyse above, the highest payoffs that a pure strategy obtains is 1 with the related pure strategy CONTINUE at both intersections. However, when randomisation is allowed, the expected payoffs can reach $4/3$ with the related behavioural strategy (CONTINUE, EXIT) = ($2/3$, $1/3$) at e_1 and e_2, and it is higher than the payoffs associated to any pure strategy.

The reason why it may be impossible to find a payoff-equivalent mixed strategy for a behavioural strategy in a game with absentmindedness is as follows. An information set with absentmindedness is executed more than once in a play of the decision problem, and then the random element in a behavioural strategy is realised independently each time when this information set is visited. On the other hand, the random element in a mixed strategy works only once and prior to the starting point of the problem. Any successive history following that information set is possible to be reached by a behavioural strategy while only part of those successive histories is able to be reached under a mixed strategy. Therefore, it is reasonable to discuss over behavioural strategies rather than mixed strategies when we analyse decision problems with imperfect recall.

Belief System and Time Consistency

If an information set is not a singleton and reachable by some strategy σ, the decision maker should form beliefs which specifying the probability of being at each history currently. Define a **belief system** as a function

μ which assigns to any history $h \in X$ a probability $\mu(h|X)$ such that $\sum_{h \in X} \mu(h|X) = 1$. It can be interpreted that when reaching information set X, the decision maker believes there is the probability of $\mu(h|X)$ that he is at history h.

PR propose the concept of consistent beliefs which is similar to the concept of consistent belief in sequential equilibrium in an extensive game with perfect recall. Assume that an information set is reachable, $\mu(h|X)$ should be equal to the proportion of times that the decision maker visits this information set and being at h in the long run if the decision maker plays the decision problem again and again while executing the strategy σ. An assessment is a pair (μ, σ). Therefore, we say an assessment (μ, σ) is **consistent (weakly consistent)** if for any information set X which is reached with positive possibility and for every $h \in X$,

$$\mu(h|X;\sigma) = p(h|\emptyset;\sigma) / \sum_{h' \in X} p(h'|\emptyset;\sigma).$$

The definition of consistency only restricts on beliefs at information sets which can be reached with positive probability. For decision problems with perfect recall, the requirement of consistency is equivalent to Bayes' formula. However, for a decision problem with absentmindedness, the decision maker could visit an information set more than once. The decision maker is said to reach a history if he either is at that history or he has passed that history. If a decision maker is stopped in an information set presenting absentmindedness, it is possible that the decision maker reaches different histories but he can never be at different histories at the time of being stopped.The sum of probability that he reach every history in an information set might be larger than one, as showed in Absentminded Driver paradox. Assume the strategy of the driver is (CONTINUE, EXIT) = $(p, 1-p)$, the probability of reaching e_1 is 1 while the probability of reaching e_2 is p, therefore the sum is $1+p$ which is larger than 1 as long as $p > 0$. The definition is similar to Bayes' formula only in notation.

Notice that the definition of a strategy requires the decision maker specifies decision rules at information sets that may not be reached during the actual play of the decision problem. We call an information set is $\sigma-$ **relevant** if it can be reached by implementing the strategy σ, i.e., $p(X|\emptyset;\sigma) > 0$. As defined in Battigalli (1997), two behavioural strategies σ and σ' are **equivalent**, i.e., $\sigma \cong \sigma'$, if they induce the same class of relevant information sets and indicate the same actions at each relevant information set. Denote $[\sigma]$ the **class** of equivalent strategies to σ. Different strategies in $[\sigma]$ behave diversely in unreachable (prevented by σ) information sets and the part of strategies restricted to those information sets can be interpreted as the behaviour induction once the decision maker deviates from his original strategy.

With a consistent belief system, we can define time consistency. A strategy σ is **time consistent** if there is a consistent assessment (μ, σ) such that for all $\sigma-$relevant information sets X and any behavioural strategy σ', such that

$$\sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|h;\sigma)u(z) \geq \sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|h;\sigma')u(z).$$

Modified Multiself Consistency and Constrained Time Consistency

We call a behavioural strategy σ is **modified multiself consistent** if there exists a consistent assessment (μ, σ) such that for every $\sigma-$relevant information set X and for every action $a \in A(X)$ for which $\sigma(h)(a) > 0$ for $h \in X$, there is no $a' \in A(X)$ such that

$$\sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|(h,a'),\sigma)u(z) > \sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|(h,a),\sigma)u(z).$$

A strategy is ex-ante optimal if, upon reaching an information set, there is no modified one-shot deviation. In decision problems with perfect recall, modified multiself consistency is equivalent to time consistency.

A behavioural strategy σ is **constrained time consistent** if there is a consistent assessment (μ, σ) such that, for all $\sigma-$relevant information sets, there is no alternative strategy $\sigma' \in [\sigma]$, σ' coincides with σ at all $\sigma-$relevant information sets except X, chooses some action $a \in A(X)$ with probability one, and yields higher expected payoffs conditional on X, i.e.,

$$\sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|h;\sigma)u(z) \geq \sum_{h \in X} \mu(h|X;\sigma) \sum_{z \in Z} p(z|(h,a);\sigma')u(z).$$

Every ex-ante optimal strategy is constrained time consistent. Battigalli (1997) also emphasises the equivalence between a constrained time consistent strategy and modified multiself equilibrium. In decision problems with perfect recall, constrained time consistency is equivalent to time consistency.

Psychological Games

To simplify the notation and prevent confusion, instead of establishing a formal framework of psychological games in this paper, we are going to define some relevant concepts and explain how to incorporate those concept into our model. We explain the framework in a game form.

A **first order belief** $b_i^1 \in B_i^1$ for player $i \in \mathbf{N}$ (utility function u_i) is a probability measure over the product of other players' behavioural strategy sets. The first belief of player i on player j's strategy is presented as $b_i^1(\sigma_j)$. Denote Σ_i the behavioural strategy set of player i, $\Sigma_{-i} = \times_{j \neq i} \Sigma_j$ and $\Sigma := \times_{i \in \mathbf{N}} \Sigma_i$. $\Delta(R)$ denotes the set of Borel probability measures on R. Then, $B_i^1 := \Delta(\Sigma_{-i})$. Although it is not mentioned in our model, a player's higher order beliefs is defined as his belief on lower order beliefs. Denote B_i^k player i's k**th order belief** set, $B_{-i}^k := \times_{j \neq i} B_j^k$ and $B^k := \times_{i \in \mathbf{N}} B_i^k$, $B := \times_{k=1}^{\infty} B^k$. Then, for $k \geq 1$,

$$B_i^{k+1} = \Delta(\Sigma_{-i} \times B_{-i}^1 \times \cdots \times B_{-i}^k).$$

A psychological Nash equilibrium is achieved when all the players' beliefs correspond to reality and no player has motivation to deviate from his equilibrium strategy while assuming other players execute their equilibrium strategies. Thus, a **psychological Nash equilibrium** of a normal form psychological game is a pair $(b, \sigma) \in (B, \Sigma)$ such that

$$b^\star(\sigma) = \sigma^\star;$$

for each $i \in \mathbf{N}$ and $\sigma_i \in \Sigma_i$,

$$u_i(b_i^\star, \sigma^\star) \geq u_i(b_i^\star, (\sigma_i, \sigma_{-i}^\star)).$$

The first condition of psychological Nash equilibrium illustrates that, for any player i, every relevant belief on his equilibrium strategy $b_j^{k,\star}(\sigma_i)$, $j \in \mathbf{N}$, coincides with $\sigma_i^\star$ in equilibrium. According to GPS, a psychological game has a psychological Nash equilibrium when the utility function $u_i : B_i \times \Sigma \to \mathbb{R}$ is continuous for each player i.

The belief in psychological games framework is different from that over information sets. The belief in psychological games shows players' expectation of other players' strategies and thoughts about others' strategy while the belief on information set which is not a singleton is the probability distribution on history that the player is at. The common point is that players form beliefs on unknown information, others' strategies or their current position.

1.6.2 Psychological Multiself Equilibrium

In this paper, we discuss finite extensive decision problems with imperfect recall (mainly, focus on absentmindedness). For a decision maker with imperfect recall, he knows the structure of the decision problem tree no matter where he makes decisions. Thus, he knows he will present imperfect recall later and which situations he will be distinguishable even if he has not reached the information set presenting imperfect recall. Once he reaches an information set with imperfect recall, he cannot distinguish the histories in the same information set by observing the surrounding environment. For the aspect of locating the decision maker himself, without any prior probability distribution or any information about his previous strategy, it is not different between being at node e_2 or e_3 in Figure 3.4 and the node e_1 or e_2 in Figure 3.1. The difference between the two situations is the way that the decision maker evaluates the decision problem at the information set presenting imperfect recall. Being at node e_1 in Figure 3.1, he knows his current behaviour not only influences the probability to each terminal nodes, but also influences his position belief of this information set. The decision maker knows he will experience the identical situation again if he moves to node e_2. However, being at node e_2 in Figure 3.4, the decision maker knows he will move out of the current information set by one-shot action. His current behaviour has no influence on his behaviour at node e_3. He would never meet the situation of reaching nodes e_2 and e_3 in the same round play of the decision problem.

A finite extensive decision problem Γ (with utility function u and n non-terminal histories in the decision tree) has a corresponding multiself psychological game form $PG(\Gamma)$. In the psychological game $PG(\Gamma)$, there are players, a planner and multiple doers. The planner is the decision maker before the decision problem processes. The planner choose a full strategy for the whole decision problem which specifies how the decision maker acts at every node. Thus, the strategy of a planner in $PG(\Gamma)$ is defined the same as the decision maker's strategy in Γ. A different doer is assigned to each history for the decision problem at execution stage. We assign a number to each node (history) in the decision problem tree. Thus, a doer's strategy is the decision maker's immediate behavioural rule at the node where that doer is located.

• The planner's behaviour strategy is $\eta = (\eta_1, \eta_2, \cdots, \eta_n)$. It assigns a probability distribution on $A(h)$ to each history h. Identical probability distribution is assigned to two histories if the two histories belong to the same information set, i.e. $\eta(h) = \eta(h')$ if $h.h' \in x \in X$.

• The strategy for doer i is φ_i and the collective strategy for doers is $\varphi = (\varphi_1, \varphi_2, \cdots, \varphi_n)$.

• The notion of utility functions. The utility function of a confident planner is U and a knowledgable planner is U^1. The utility function for doer i is V_i.

• Belief system. There are two types of belief in the model. One is similar to that developed in games with imperfect information. A decision maker deduces which history is at currently by the belief μ. An assessment (μ, σ) is consistent if the belief is defined by strategy σ following Bayes' rule. We call it **position belief**. The other belief system is similar to what is described in psychological games framework. The different selves form belief $b \in B$ on other selves' behaviours and beliefs. Denote $b^i_j(\sigma)$ the ith belief of doer j on the strategy σ. Particularly, denote b^1_{pl} the the first belief of the planner. We call this part of belief **behaviour belief**.

A doer forms the position belief using the information of the planner's strategy if the doer is located in an information set with imperfect recall. The reason that we use a similar definition of belief in games with imperfect information is that the selves are imperfect informed of other selves' behaviours or information if we translate the decision problem as a multiself model. The planner evaluates different strategies according to his different behaviour beliefs about doers. In the model, only the planner's first order behaviour belief on doers' actions is involved. The planner's first order behaviour belief contains the planner's belief about how the doers behave.

A doer chooses the optimal strategy based on the utility function. Since each doer in $PG(\Gamma)$ is the decision maker in Γ at each node in the stage of execution and the decision maker is reevaluating the play in that stage, the utility function of a doer is the conditional expected payoff function of the decision maker at the node where the doer is. Furthermore, since we assume the decision maker's control power is restricted to his current behaviour, each doer's strategy is a one-shot move. Once a doer is located in an information set with imperfect recall, the doer could deduce his precise position by the strategy of the planner. The planner moves first and the planner's chosen strategy is the known information of doers. Then, the utility function of doer i (decision maker at $h_i \in X_j \in \mathscr{X}$) $V_i : \Sigma(h_i) \to \Re$, $i \in \{1, \cdots, n\}$ is

$$V_i(\varphi_i) = V_i(\varphi_i|\eta, \varphi_{-i}) = Eu(\varphi_i|h_i \in X_j; \varphi_{-i}, \eta) = \sum_{h \in X_j} \mu(h|X_j; \eta) \sum_{z \in Z} p(z|(h, \varphi_i); \varphi_{-i}) u(z),$$

In the utility function, $\mu(h|\eta, X_j)$ is the relative probability of being at history h in the information set X_j which is deduced by the doer i from the planner's strategy η. If there exists a consistent assessment (μ, η), the position belief of the doer is

$$\mu(h|X_j; \eta) = \frac{p(h|\emptyset; \eta)}{\sum_{h' \in X_j} p(h'|\emptyset; \eta).}$$

We assume the doers form their position beliefs by the planner's strategy. The planner acts before the decision problem starts, and doers know the planner's strategy when they are going to act. The reason why the doers deduce his current position by the planner's strategy is that it is the only evidence that the doers hold about his current position since the latest strategy cannot be delivered to the decision maker. From the viewpoint of a doer, he is not sure about other doers' strategies. The doer does not know either whether other doers reconsider or what strategy they deviate once they reconsider. Besides that, the motivation of our approach is to examine whether the ex-ante strategy is conditionally optimal. Then, the belief at a non-singleton information set should take advantage of the ex-ante strategy. Thus, if there exists a belief system consistent with the planner's strategy, a doer forms their position belief which follows the Bayesian rule with the strategy of the planner.

In fact, the nature of absentmindedness presents the cycle problem of “strategy-belief-strategy". The cycle is explained as the following. The strategy at an information set with absentmindedness determines the decision maker’s position belief, and at the same time, the decision maker’s position belief influences the decision maker’s choice of strategy. However, the doers form position beliefs using the planner’s strategy in our approach. Thus, the approach separates the process of belief forming and strategy choosing which makes the approach reasonable.

A confident planner believes doers will follow the strategy he chooses. From the multiself aspect, the planner is confident that doers will listen to his order. From one-self aspect, the decision maker is confident that his ex-ante strategy will still be optimal when the decision problem is executed. Thus, the confident planner’s utility function is the ex-ante expected payoff function of the decision maker in Γ. The utility function of a confident planner $U : \Sigma \to \Re$ is

$$U(\eta) = Eu(\eta) = \sum_{z \in Z} p(z|\emptyset; \eta) u(z).$$

We can see that there is no first order behaviour belief of the confident planner in the utility function. Since the decision maker believes the doers implement the behavioural rules described in the planner’s strategy, it implies the first order behaviour belief of the confident planner is $b^1_{pl}(\varphi) = \eta$.

A knowledgable planner notices doers will reevaluate the decision problem based on their current nodes and it is possible for doers to have multiple one-shot deviations from the planner’s strategy. The multiple one-shot deviations indicate the decision maker might deviate multiple times but only change the immediate behaviour in each deviation. There is no collaboration among doers due to imperfect communication. The planner knows how doers evaluate the decision problem and believe doers are rational. Thus, the knowledgable planner’s first order behaviour belief is that the doers’ strategies are the best responses to the planner’s strategy. The knowledgable planner would like to help the doers maximise their payoffs. Thus, the knowledgable planner should choose a strategy which makes the doers believe they are in a larger probability in the situation that they can achieve high payoffs and in a smaller probability in the situation that they acquire low payoffs. In other words, the planner would like to help the doers seek relative profits and avoid relative losses since the planner knows he and all the doers is a team. The knowledgable planner is knowledgable compared to the confident planner in the aspect of how doers consider the decision problem. The knowledgable planner’s utility function

$U^1 : (B^1_{pl} \times \Sigma) \rightarrow \Re$, where $B^1_{pl} := \Delta(\times_{i\in\{1,\cdots,n\}} \Sigma(h_i))$, is

$$U^1(b^1_{pl}(\varphi),\eta) = \sum_{X\in\mathscr{X}} Eu(\eta|X;b^1_{pl}(\varphi)) = \sum_{X\in\mathscr{X}} \left[\sum_{h\in X} \mu(h|X;\eta) \sum_{z\in Z} p(z|h;b^1_{pl}(\varphi))u(z)\right],$$

with the planner's first order behaviour belief on doers' strategies, $b^1_{pl}(\varphi) \in BR_\varphi(\eta)$ and for any η,

$$BR_{\varphi_i}(\eta,\varphi_{-i}) = \{\hat{\varphi}_i \in \Sigma(h_i)|V_i(\hat{\varphi}_i|\eta,\varphi_{-i}) \geq V_i(\varphi_i|\eta,\varphi_{-i}), \forall\varphi_i \in \Sigma(h_i)\},$$

$$BR_\varphi(\eta) = \{\varphi \in \times_i \Sigma(h_i)|\varphi_i \in BR_{\varphi_i}(\eta,\varphi_{-i})\}.$$

The set $BR_\varphi(\eta)$ is the fixed point set of the correspondences $BR_{\varphi_i}(\eta,\varphi_{-i}), i \in \{1,\cdots,n\}$, where BR_{φ_i} represents the doer i's best response to other doers's collective strategy φ_{-i} and the planner's strategy η. Notice that the utility function of a knowledgable planner is identical to the utility function of doers. The difference between them is that the planner chooses a strategy to establish the doers' belief system.

Definition 1.6.1 *For a finite decision problem Γ with the utility function u, we could present the decision problem Γ as a **psychological multiself game** $PG(\Gamma)$ as follows,*

• *A distinct doer is located at every node and his strategy is a one-shot move. The utility function of the doer i $V_i : \Sigma(h_i) \rightarrow \Re$, $i \in \{1,\cdots,n\}$ is*

$$V_i(\varphi_i) = \sum_{h\in X_j} \mu(h|X_j;\eta) \sum_{z\in Z} p(z|(h,\varphi_i);\varphi_{-i})u(z).$$

• *A planner acts before the decision problem executes and the planner chooses a strategy indicating the decision maker's behaviours at every node. The different type of planner hold different beliefs about the doers strategies.*

• *The utility function of a confident planner $U : \Sigma \rightarrow \Re$ is*

$$U(\eta) = \sum_{z\in Z} p(z|\emptyset;\eta)u(z),$$

with the first order behaviour belief

$$b^1_{pl}(\varphi) = \eta.$$

• *The utility function of a knowledgable planner* $U^1 : (B^1_{pl} \times \Sigma) \to \Re$ *is*

$$U^1(b^1_{pl}(\varphi), \eta) = \sum_{X \in \mathscr{X}} \left[\sum_{h \in X} \mu(h|X;\eta) \sum_{z \in Z} p(z|h; b^1_{pl}(\varphi)) u(z) \right],$$

with the first order belief

$$b^1_{pl}(\varphi) \in BR_\varphi(\eta),$$

where

$$BR_{\varphi_i}(\eta, \varphi_{-i}) = \{\hat{\varphi}_i \in \Sigma(h_i) | V_i(\hat{\varphi}_i|\eta, \varphi_{-i}) \geq V_i(\varphi_i|\eta, \varphi_{-i}), \forall \varphi_i \in \Sigma(h_i)\},$$

$$BR_\varphi(\eta) = \{\varphi \in \times_i \Sigma(h_i) | \varphi_i \in BR_{\varphi_i}(\eta, \varphi_{-i})\}.$$

We can see from the definition of a psychological multiself game form with a confident planner and with a knowledgable planner that, the difference between the two descriptions is that the planner's behaviour belief about doers. The behaviour belief is the planner's internal information which cannot be observed by others. The knowledgable planner presents how the decision maker deals with the choice of an ex-ante strategy once he realises how he acts in the stage of execution and expects possible deviations from his current chosen strategy. The structures of a confident planner and a knowledgable planner make a complete psychological multiself game. Of course, the planner could form other behaviour beliefs about the doers when he realises potential deviations during execution. The scenario behind the different behaviour beliefs is different from what we discuss in this paper. As long as it is a reasonable description, it can also be included in the psychological multiself game form of a decision problem with imperfect recall.

Notice that the belief of the planner in the definition of a PMEK is presented as a first order belief. It is because we describe the situation as the planner knows the doers know the strategy that he chooses before the decision problem starts. Thus, the ex-ante chosen strategy is certain information and common knowledge of the doers. Then, the first order belief of the planner can be phrased as the planner believes doers will choose a strategy which belongs to their best responses. Alternatively, we could also define the belief in a PMEK as a second belief. In such a case, the planner does not know the doers know his chosen strategy. Instead, he thinks the doers will implement one of their best responses to whatever strategy they believe the planner has chosen. The belief system for such a knowledgable planner includes an implicit the first order belief and the second order belief. The first order belief is that doers will take the ex-ante strategy they believe the planner has chosen as the evidence to form their belief on the probability distribution of being at each history. The second order belief is the planner's belief of the doers' beliefs of his chosen strategy. In fact, the second assumption is little reasonable since the existence of the planner becomes meaningless if the doers do not know the strategy that the planner chooses before the decision problem starts.

The definition 1.6.1 describes how we transfer a finite decision problem Γ into a psychological multiself game $PG(\Gamma)$. Since the framework of a psychological multiself game is established, we are going to discuss the equilibrium issue of the game.

The definition of psychological games indicates that, if a psychological Nash equilibrium exists, the equilibrium strategy should maximise the utility function while the equilibrium belief should be coincide with the equilibrium strategy. Thus, the psychological multiself equilibrium is defined as follows.

Definition 1.6.2 *We say that a decision problem Γ which can be interpreted as a psychological multiself game form $PG(\Gamma)$ has a* ***psychological multiself equilibrium with a confident planner (PMEC)*** $(b^1_{pl}, \eta^\star, \varphi^\star) \in (B^1_{pl}, \Sigma, \times_{i\in\{1,\cdots,n\}} \Sigma(h_i))$ *if*
1) there exists a position belief μ consistent with $\eta^\star$;
2) for any $\eta \in \Sigma, \rho \in \Sigma(h_i)$, we have

$$b^1_{pl}(\varphi) = \eta^\star = \varphi^\star,$$

$$U(\eta^\star) \geq U(\eta),$$

$$V_i(\varphi^\star_i | \eta^\star, \varphi^\star_{-i}) \geq V_i(\rho | \eta^\star, \varphi^\star_{-i});$$

3) for any collective strategy of the doers $\varphi \neq \varphi^\star, \varphi \in \bigcup_{h\in H} \Sigma(h)$, one of the following three conditions must be satisfied.
• *there exists a doer i and $\rho \in \Sigma(h_i)$ such that*

$$V_i(\rho | \eta^\star, \varphi_{-i}) > V_i(\varphi_i | \eta^\star, \varphi_{-i});$$

• *for any doer* $i \in \{1,...,n\}$,

$$V_i(\varphi_i|\eta^\star,\varphi_{-i}) \leq V_i(\varphi_i^\star|\eta^\star,\varphi_{-i}^\star);$$

• *there exists doer i and k in the same information set, such that*

$$\varphi_i \neq \varphi_k.$$

The decision problem has a ***psychological multiself equilibrium with a knowledgable planner (PMEK)*** $(b_{pl}^1,\eta^\star,\varphi^\star) \in (B_{pl}^1,\Sigma,\times_{i\in\{1,\cdots,n\}}\Sigma(h_i))$ *if*
1) there exists a position belief μ *consistent with* $\eta^\star$*;*
2) for any $\eta \in \Sigma$*, we have*

$$b_{pl}^1(\varphi) = \varphi^\star \in \Sigma,$$

$$U^1(b_{pl}^1(\varphi),\eta^\star) \geq U^1(b_{pl}^1(\varphi),\eta);$$

$$V_i(\varphi_i^\star|\eta^\star,\varphi_{-i}^\star) \geq V_i(\rho|\eta^\star,\varphi_{-i}^\star);$$

In equilibrium, the planner chooses an optimal strategy according to his different beliefs on doers' behaviours. A confident planner believes doers will follow his chosen strategy. Thus, the confident planner in $PG(\Gamma)$ behaves the same as a decision maker in Γ, and chooses an ex-ante optimal strategy in Γ. His behaviour belief of doers' collective strategy should be identical to the doers' actual moves, and is the same as his chosen optimal strategy. Since a PMEC requires the equilibrium collective strategy of doers to be identical to an equilibrium strategy of the planner, "time consistency" is required in the existence of an equilibrium in the sense that we say a strategy is time consistent if the collection of doers' behaviours is the same to the planner's strategy. A knowledgable planner knows the fact that the doers will not directly follow his chosen strategy but use it as evidence to deduce their current position in the decision problem tree. The knowledgable planner would like to evaluate the same way as the doers and help them seek for more profits and avoid fewer losses (the profits and losses are the relative concepts here). He knows the doers will evaluate the decision problem based on their current expected payoffs and choose an immediate strategy to achieve the highest conditional expected payoffs. Thus, the knowledgable planner believes the doers will play the strategies from their best responses to his chosen strategy and each other's strategy. In equilibrium, doers behave the same way as the knowledgable planner believes. As for the doers, since the planner moves before the decision problem executes, the planner's strategy is known to the doers and it is common knowledge among doers. If a doer is located in an information set with imperfect recall, he will take advantage of the planner' strategy to form his position belief which is the probability distribution of being at each history in that information set. In equilibrium, the doers play strategies from their best responses to each other.

The definition of a PMEC and a PMEK illustrates the following several points. The first condition in definition 1.6.2 indicates that the doers' position belief about the probability distribution of his current node in the information set should be consistent with the planner's strategy in equilibrium. The second condition describes the requirements for a psychological equilibrium. The planner chooses an optimal strategy with his behaviour belief about doers. Each doer plays one of the best responses to the others' strategies with their common knowledge of the planner's strategy. Thus, if a strategy belongs to a PMEC, no one-shot deviation occurs when the decision problem is running. Since our approach analyses multiple reconsiderations during execution, the strategy from a PMEC must satisfy the condition if the reconsideration occurs once, i.e., modified multiself consistency, because reconsideration once is a special case of multiple reconsiderations. The third condition presents the difference between a PMEC and a modified multiself equilibrium. It explains the reason why there are no deviations among doers when executing the equilibrium collective strategy of doers σ. For any other strategy σ', either the strategy is not an equilibrium among the doers, then there is always at least a doer who deviates from this strategy σ'(the first point of 3) in definition 1.6.2). Or, no doer could achieve high expected payoffs by deviations from the equilibrium collective strategy σ(the second point of 3) in definition 1.6.2). Every doer has no motivation to deviate even though he assumes the deviations from other doers. In this situation, the strategy σ' might be an equilibrium collective strategy.

The third point of 3) in definition 1.6.2 presents the internal consistency. It implies each doer should assume the same behaviour of the other doers in the same information set. Notice that, there might be another asymmetry equilibria among doers, which makes the doers in the same information set choose different strategies in equilibrium. Back to the psychological multiself game with a confident planner of the example of absentminded driver paradox, besides the behavioural strategy of $(2/3, 2/3)$, there are two equilibrium collective strategy $(1,0)$, $(0,1)$. If such an equilibrium among doers exists, it indicates that the decision maker who chooses a behavioural strategy ex-ante could deviate to other collective strategies of doers if the decision maker is allowed to reconsider more than once in the same information set and regardless of the belief of the planner. However, the fact that the planner's belief on doers' behaviours must be a behavioural strategy. For an asymmetry equilibrium collective strategy of doers, each doer chooses the behavioural rule which is the best response to each other but the collection of their strategy is not a part of any PMEC since a PMEC requires the collection of doers' strategy is identical to the chosen strategy of the planner.

Besides that, in real life, an individual always receives the same issue in the same way if there is no new information is learned between two issues occurs, especially the decision maker makes successive decisions. The internal consistency of people is an underlying assumption of our approach. For example, in the absentminded driver paradox, the decision maker cannot observe the actions he chooses at other histories in the same information set. Thus, when he chooses his optimal strategy, he needs to assume the behaviours at other histories, although other nodes' behaviours are presented as known in the analysis. Then, assume nodes h_1 and h_2 are in the same information set. Since the decision maker is assumed to totally forget his currently accurate location in the decision problem tree, the decision problem that he faces at the two nodes should be identical. If the decision maker at h_1 assumes his action at h_2 is A, when he moves to h_2, his assumption of the action at h_1 should also be A. Then, they make their optimal decisions according to their assumptions. However, if we do not suppose the internal consistency of individuals, the asymmetry equilibrium could appear. The decision maker at h_1 could assume his action at h_2 is A and he at h_2 assumes his action at h_1 is A', even if he faces identical decision problems at both nodes.

However, the internal consistency is less reasonable if we apply the approach to the manager-agents model. The internal consistency in the manager-agents model implies that the manager thinks the agents assume the same others' behaviours when they meet the identical decision problem. However, the perception of the environment is hard to be identical among individuals. Thus, we require the equilibrium in manager-agents to be identical and symmetry because the manager cannot distinguish the different nodes in the same information set and believes doers should follow his order and execute his chosen strategy for the decision problem.

The third point of 3) in definition 1.6.2 indicates that, even if a strategy σ', which is not from a PMEC, is an equilibrium collective strategy and all doers reach higher expected payoffs by deviations, the doers in the same information set assume different behaviours of the others in equilibrium. It contradicts our underlying requirement for internal consistency of recognition.

In our approach, there is one type of doer and two types of planner. The planner assumes different behaviours of the doers, but the doers are assumed to behave the same way with both beliefs of the planner. For a knowledgable planner, he will assume the doers in the same information set meet identical situations. Assume doer i and doer j are located in the same information set. The doer i assumes the doer j's strategy is A when the doer i reconsiders the decision problem. Then, we require that the doer j assumes the doer i's strategy is A, too. As shown in Figure 3.3, We say that a PMEC exists if $b^1_{pl}(\varphi) = \eta = \varphi$. Even though the collective equilibrium strategy of doers is the same as the planner's equilibrium strategy, the doers reconsider the decision problem and choose their optimal strategies instead of just following the planner's strategy. It seems to be contradicted by the definition of psychological Nash equilibrium. However, what the planner can observe is the doers' behaviours but not how they choose the strategies. How doers behave is doers' internal information which cannot be seen directly.

The psychological multiself equilibrium has been defined. Next, we are going to prove the existence of a PMEC and PMEK.

Proposition 1.6.1 *For a decision problem Γ and its psychological multiself form $PG(\Gamma)$, a PMEC and PMEK exist if the utility function $u : \Sigma \to \Re$ of Γ is finite.*

Before explaining the existence of an equilibrium, we prove the following proposition which indicates the equivalence of an optimal strategy in Γ and a PMEC in $PG(\Gamma)$.

Lemma 1.6.1 *The induced utility functions for the confident planner $U : \Sigma \to \Re$, for the knowledgable planner $U^1 : (B^1_{pl} \times \Sigma) \to \Re$ where $B^1_{pl} := \Delta(\times_{i\in\{1,\cdots,n\}} \Sigma(h_i))$, for the doer i $V_i : \Sigma(h_i) \to \Re$, $i \in \{1,\cdots,n\}$, for any (η, φ_{-i}) and for the conditional expected payoff function at history h_i $V^1_i : \Sigma \times (\times_i \Sigma(h_i)) \to \Re$ are continuous* *if the utility function $u : \Sigma \to \Re$ is finite.*

Proposition 1.6.2 *For a finite decision problem Γ with utility function u, a behaviour strategy σ is ex-ante optimal in Γ if and only if $(b^1_{pl}, \eta, \varphi) = (\sigma, \sigma, \sigma)$ is a PMEC in $PG(\Gamma)$.*

Proof of Proposition 1.6.2. $\Leftarrow$ First, in $PG(\Gamma)$, for a PMEC $(b^{1,\star}_{pl}, \eta^\star, \varphi^\star)$, we have $b^{1,\star}_{pl} = \eta^\star = \varphi^\star$. The equilibrium strategy $\eta^\star$ of the confident planner satisfies

$$\eta^\star \in \underset{\eta}{argmax}\, U(\eta) = \underset{\eta}{argmax}\, Eu(\eta) = \underset{\eta}{argmax} \left(\sum_{z \in Z} p(z|\emptyset;\eta) u(z) \right).$$

Then, the strategy $\eta^\star$ is an ex-ante optimal strategy for decision problem Γ.

$\Rightarrow$ If a strategy σ is an ex-ante optimal strategy for decision problem Γ, we need to prove the existence of a PMEC $(b^{1,\star}_{pl}, \eta^\star, \varphi^\star) = (\sigma, \sigma, \sigma)$ for $PG(\Gamma)$. Notice that the behaviours of doers assigned to later histories are not influenced by former doers if they are not in the same information set and all the doers have identical knowledge about the confident planner's strategy. The process of proof is similar to generalised backward induction procedure in games with imperfect information.

1. For each of the final information set X of $PG(\Gamma)$ that can be reached by σ, doers in that information set form an Nash equilibrium based on their knowledge of the confident planner's strategy σ. We need to prove that there exists an collective equilibrium strategy of doers which is the same as $\sigma(X)$.
2. We choose the collective equilibrium strategy of doers in this final information set which is the same as $\sigma(X)$, then derive the reduced extensive decision problem in which these final information sets are replaced by the payoffs that result in these information sets when doers use their equilibrium strategies. Also derive corresponding psychological multiself game form of the reduced extensive decision problem.
3. Repeat step 1 and 2. Continue the procedure until the moves at histories which belongs to σ-relevant information sets are determined. For the doers at unreachable information sets by σ, we define their behavioural rules to be the same as σ. Then, the collection of moves at various histories constitutes a collective equilibrium

strategy $\varphi^\star$ for doers, and $\varphi^\star = \sigma$.

Notice that multiple equilibria may encounter in some steps of this process, we only need to prove there exists one collective equilibrium strategy of doers which is identical to the confident planner's strategy, thus the optimal strategy σ in decision problem Γ.

Now, we prove that for each final information set X of $PG(\Gamma)$, there is a collective equilibrium strategy of doers in the information set X such that $\varphi^\star(X) = \sigma(X)$. First, it is easy to see $\varphi^\star(X) = \sigma(X)$ if the final information set X is a singleton. When the final information set X is not a singleton, there must be a non-singleton information set such that any path goes through it does not pass any non-singleton information set till terminal histories, we start from that information set.

Assume histories $h_{i_1}, \cdots, h_{i_k} \in X$, then for each h_{i_j}, we only need to prove the optimal strategy for doer i_j is $\sigma(X)$ when the other doers in this information set choose $\sigma(X)$. We know that each doer at history h, who has subhistories in X, chooses $\sigma(h)$ in equilibrium. Thus, we need to prove

$$\sigma(X) = \underset{\varphi_{i_j}}{argmax} V_{i_j}(\varphi_{i_j}|\sigma, \varphi_{-i_j}) = \underset{\varphi_{i_j}}{argmax}\left(\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|(h,\varphi_{i_j});\varphi_{-i_j})u(z)\right).$$

In the equation, φ_{-i_j} is a collective strategy that the doers at histories in X and those who have subhistories in X follow the same behavioural rule as σ. We know that the choice of optimal strategy for doer i_j is irrelevant to behaviours at previous histories of information set X. Then, the problem is the same as to prove

$$\sigma(X) = \underset{\varphi_{i_j}}{argmax}\left(\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|(h,\varphi_{i_j}),\sigma)u(z)\right).$$

It is easy to see that the problem becomes to prove an optimal strategy is modified multiself consistent. It is proved in PR. Thus, the doers at a non-singleton information set also follow σ in equilibrium. By backward induction, we know that for any optimal strategy σ, there is a psychological equilibrium $(b_{pl}^{1,\star}, \eta^\star, \varphi^\star) = (\sigma, \sigma, \sigma)$ for $PG(\Gamma)$.

Now, we only need to prove that other collective strategy of doers must satisfy one of the three conditions in the definition of a PMEC (see Definition 3) in 1.6.2). Assume there is a collective strategy of doers φ' which does not satisfy any of the three conditions in the definition of a PMEC. Then, φ' must be a behavioural strategy, $\varphi' \in \Sigma$, and it is an equilibrium among doers. In other words, for any doer i and $\rho \in \Sigma(h_i)$,

$$V_i(\varphi_i'|\sigma, \varphi_{-i}') \geq V_i(\rho|\sigma, \varphi_{-i}').$$

More importantly, there exists a doer $i \in \{1, ..., n\}$, $h_i \in X_j$ such that

$$V_i(\varphi'|\sigma, \varphi_{-i}') > V_i(\sigma_i|\sigma, \sigma_{-i}).$$

It is equivalent to

$$\sum_{h \in X_j} \mu(h|X_j;\sigma) \sum_{z \in Z} p(z|h,\varphi')u(z) > \sum_{h \in X_j} \mu(h|X_j;\sigma) \sum_{z \in Z} p(z|h,\sigma)u(z),$$

$$\sum_{h \in X_j} p(h|\emptyset;\sigma) \sum_{z \in Z} p(z|h,\varphi')u(z) > \sum_{h \in X_j} p(h|\emptyset;\sigma) \sum_{z \in Z} p(z|h,\sigma)u(z). \tag{1.2}$$

Denote σ_1 the behavioural strategy such that

1. the decision maker implements φ' at all histories in information set X_j and reaches the information set $X_{j_1}, ..., X_{j_k}$;
2. the decision maker implements φ' at all histories in information sets $X_{j_1}, ..., X_{j_k}$ and reaches some more information sets;
3. repeat the process 2 until the decision maker reaches terminal histories,
4. the decision maker implements the strategy σ at other histories.

Also denote σ_2 the behavioural strategy such that

1. the decision maker implements φ' at all histories in information set X_j and reaches the history $X_{i_1}, ..., X_{i_t}$;
2. the decision maker implements φ' at a history h, if all the other histories in the same information set with h can be reached in process 1; 3. repeat the process 2 until the decision maker reaches terminal histories,
4. the decision maker implements the strategy σ at other historiesin the decision tree.

Then, the value of term

$$\sum_{h \in X_j} p(h|\emptyset;\sigma) \sum_{z \in Z} p(z|h,\varphi')u(z)$$

must be in between the value of the two terms

$$CEP1 = \sum_{h \in X_j} p(h|\emptyset;\sigma_1) \sum_{z \in Z} p(z|h,\sigma_1)u(z)$$

and

$$CEP2 = \sum_{h \in X_j} p(h|\emptyset;\sigma_2) \sum_{z \in Z} p(z|h,\sigma_2)u(z).$$

The strategy σ is ex-ante optimal. For any behavioural strategy $\bar{\sigma}$, we have

$$\sum_{z \in Z} p(z|\emptyset;\sigma)u(z) \geq \sum_{z \in Z} p(z|\emptyset;\bar{\sigma})u(z).$$

Thus, the term

$$\sum_{h \in X_j} p(h|\emptyset;\sigma) \sum_{z \in Z} p(z|h,\sigma)u(z)$$

must be larger than or equal to $CEP1$ and $CEP2$. It is contradicted to the inequation 1.2. Thus, such a collective strategy of doers φ' does not exist. Q.E.D.

Proof of Proposition 1.6.1. From Proposition 1.6.2, we know an ex-ante optimal strategy must be a psychological multiself equilibrium with a confident planner, then the existence of such an equilibrium is proved. We are going to prove the existence of an equilibrium with a knowledgable planner.

No matter for doer i or the knowledgable planner, their strategy set $\Sigma(h_i)$ and Σ is a compact set since the decision problem Γ is finite.

For doer i, we need to prove for any planner's strategy η and other doers' strategies φ_{-i}, there is a fixed point for its best response correspondence BR_{φ_i}. By Kakutani fixed point theorem, we only need to prove that BR_{φ_i} is upper hemicontinuous, $BR_{\varphi_i}(\eta,\varphi_{-i})$ is non-empty, compact and convex for all (η,φ_{-i}).

In fact, $V_i(\varphi_i)$ is a linear function with respect to φ_i. The domain of φ_i, $\Sigma(h_i)$, is compact. It is obviously that $BR_{\varphi_i}(\eta, \varphi_{-i})$ is non-empty, compact and convex for all (η, φ_{-i}). $V_i^1(\eta, \varphi)$ is continuous on $\Sigma \times (\times_i \Sigma(h_i))$. By Maximum Theorem, BR_{φ_i} is upper hemicontinuous and compacted valued.

Thus, for any η, $BR_{\varphi}(\eta)$ is a non-empty correspondence. For the knowledgable planner, he knows how doers response to his strategy in equilibrium. Then, his optimality problem is to choose a strategy that maximises $U^1(b^1_{pl}(\varphi), \eta)$, for any $b^1_{pl}(\varphi) \in BR_{\varphi}(\eta)$. Then, U^1 is a continuous function of η and the domain Γ is compact. The maximum and maximiser of U^1 must exist.

Thus, an psychological multiself equilibrium with a knowledgable planner exists. Q.E.D.

In decision problems with perfect recall, the collective equilibrium strategy of a confident planner is the same as that of a knowledgable planner. Every information set is singleton if there is no chance player. No doer needs to deduce his position through planner's strategy. Thus, a psychological multiself equilibrium with a knowledgable planner is (σ, η, σ), σ is the optimal strategy for that decision problem and η is a strategy which can make the decision maker reach every σ-relevant information set.

1.6.3 Multiple Multiselves Consistency

The motivation of our psychological multiself approach is that the decision maker might deviate from the ex-ante optimal strategy when he reevaluates the decision problem during the execution if the imperfect recall (especially absentmindedness) presents. In other words, an ex-ante optimal strategy might not be time consistent. A strategy is time consistent indicates the strategy is conditionally optimal during execution. Then, we would like to explore whether a strategy is conditionally optimal if restricting the decision maker's control power to the behaviour assigned at the node where he reconsiders. In fact, the modified multiself consistency in PR illustrates an ex-ante strategy is conditionally optimal if the decision maker is allowed to have one-shot deviation. However, the nature structure of an information set with absentmindedness makes the assumption that the decision maker only reconsiders once during one round play of the decision problem unreasonable. Thus, we would like to see whether an ex-ante optimal strategy is conditionally optimal once the assumption is relaxed to allow the decision maker reconsiders more than once. Even if a strategy is conditionally optimal based on at most one-shot deviation, it might not maintain its conditional optimality if allowing multiple one-shot deviations. Furthermore, once the decision maker realises the probability of deviations during execution, how the decision maker chooses the ex-ante strategy is what we would like to analyse.

From the definition of PMEC, if a strategy belongs to a PMEC, the planner's strategy and the collective strategy of doers are identical to that strategy. In other words, the ex-ante chosen strategy is the same as the conditionally optimal strategy if the conditional optimality is examined by allowing multiple one-shot deviations. The equilibrium does not include time consistency as a requirement. However, a strategy belonging to a PMEC is time consistent if we allow the decision maker chooses one-shot deviations multiple times.

Definition 1.6.3 *A strategy σ is* ***multiple multiselves consistent*** *if there exists a consistent assessment (μ,σ), for any h which is reachable by implementing the strategy σ, $\forall\rho\in\Sigma(h)$, we have*

$$\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|h,\sigma)u(z)\geq\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|(h,\rho),\sigma)u(z);$$

and there is not a behavioural strategy $\sigma'\neq\sigma$ such that
1) $\forall h\in H$, $\forall\rho\in\Sigma(h)$,

$$\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|h,\sigma')u(z)\geq\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|(h,\rho),\sigma')u(z);$$

2) $\exists h\in H$,

$$\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|h,\sigma')u(z)\geq\sum_{h\in X}\mu(h|X;\sigma)\sum_{z\in Z}p(z|h,\sigma)u(z).$$

where the strict greater-than character is established at least in one inequation (there exists at least one $h\in H$ which makes the strict greater-than character established).

The second condition in the definition of multiple multiselves consistency indicates that, unless another equilibrium among doers brings at least one doer more expected payoffs, otherwise the doers do not deviate to the different strategy. Thus, if a strategy is multiple multiselves consistent, it is a stable equilibrium among doers under the premise that the decision maker uses the consistent strategy as the evidence to deduce his current position when meeting imperfect recall. The stability is presented as the fact that the decision maker during execution will not deviate to other doers' equilibrium strategy if he is executing the multiple multiselves consistent strategy.

We could see from Definition 1.6.3 that the freedom of deviation of multiple multiselves consistency is between time consistency and modified multiself consistency. The common aspect between multiple multiselves consistency and time consistency is that both of them examine whether a strategy σ is conditionally optimal if the decision maker can choose a new strategy which is different from σ at multiple nodes. The common aspect between multiple multiselves consistency and modified multiself consistency is that both of them examine conditional optimality based on allowing one-shot deviation. If the strategy σ' in Definition 1.6.3 is σ, multiple multiselves consistency and modified multiself consistency are the same. In other words, modified multiself consistency is a special case of multiple multiselves consistency. Thus, a strategy is multiple multiselves consistent if it is time consistent. A strategy is modified time consistent if it is multiple multiselves consistent.

If a strategy is multiple multiselves consistent, it indicates the decision maker cannot successfully deviate to a different strategy if firstly, he is allowed to have one-shot deviations and the reconsideration occurs at each history; and the decision maker uses the consistent strategy as the evidence to deduce his current position when he meets imperfect recall. If the decision maker is executing a strategy which is not multiple multiselves consistent, there always exists at least one time of reconsideration at which time the decision maker would like to deviate from his current behavioural rule to a more beneficial one. Assume the reconsideration which makes the current behavioural rule not conditionally optimal occurs at node h. Then, if the decision maker executes the revised strategy, he should assume he executes the revised behavioural rule at node h when he reconsiders at other nodes. If the revised strategy is still not multiple multiselves consistent, there must be at least one round of reconsideration which occurs at another history such that the decision maker can deviate to a new behavioural rule than that described in the revised strategy. Thus, when a decision maker executes a multiple multiselves consistent strategy, it is impossible to deviate to a different strategy during execution even if the decision maker assumes deviations at other histories. The available deviation of the decision maker during execution is assumed to be any behavioural strategy rather than any collective strategy of every non-terminal node. It indicates there is no deviation which assigns the same behavioural rule to the histories in the same information set, if the decision maker executes a multiple multiselves consistent strategy.

There is an underlying requirement for multiple time consistency that, if a decision maker reconsiders at a node, he must also reconsider at other nodes in the same information set as that node, and the deviations at all nodes in an information set must be identical in the assumption of the decision maker's reconsideration. It is because the multiple multiselves consistency evaluates deviation among behavioural strategies in Σ. Otherwise, the identical deviations of doers in the same information set cannot be realised if there exists a doer in that information set who does not reconsider.

Proposition 1.6.3 *A strategy from a PMEC is multiple multiselves consistent.*

Corollary 1.6.1 *An ex-ante optimal strategy is multiple multiself consistent.*

Proposition 1.6.3 and the corollary 1.6.1 are easy to be proved by the definition of a PMEC and multiple multiselves consistency.

We would like to explain the corollary 1.6.1 back to the decision problem Γ. Assume a decision maker firstly choose a strategy before the decision problem starts, and he could examine the conditional optimality of the chosen strategy at any node which is reachable by the chosen strategy multiple times. At a node, the decision maker evaluates all available immediate actions to examine whether the behavioural rule in the ex-ante strategy is conditionally optimal. When he does the reconsideration, he assumes he also reconsiders and might deviate at other nodes. Thus, in his conditional expected payoff function, the behaviours at other nodes are not necessarily identical to the ex-ante chosen strategy. Once the decision maker moves to an information set presenting imperfect recall, he would form his position belief using his ex-ante chosen strategy. In equilibrium, the decision maker chooses an ex-ante optimal strategy before the decision problem runs and the collection of the decision maker's one-shot moves at different nodes are identical to the ex-ante optimal strategy. Then, we could interpret the equilibrium in this way. In a decision problem with imperfect recall, if a decision maker chooses an ex-ante optimal strategy before the decision problem starts, he will not deviate from this strategy no matter how many times the decision maker reevaluates the strategy and considers one-shot deviations.

If comparing the definition of PMEC and the definition of multiple multiselves consistency, we could find that the definition of PMEC requires more strictly than the definition of multiple multiselves consistency. The definition of PMEC requires an equilibrium strategy should achieve more payoffs than any collective strategy of the doers, while the definition of multiselves consistency only requires the consistent strategy is more beneficial than any behavioural strategy of the decision problem. The collective strategy of the doers is identical to a behavioural strategy if the decision maker has perfect recall. However, when imperfect recall presents, a general collective strategy of doers could assign different behavioural rules at histories in the same information set. The modified multiself consistency requires the consistent strategy should obtain more expected payoffs than any pure behaviour available at each history. Compared, our multiple multiselves consistency requires a consistent strategy should be more beneficial than any behavioural rule assigned at each history. It is enough to evaluate every pure action if the decision problem does not present absentmindedness. However, as the absentminded driver paradox shows, there might be a behavioural strategy that obtains more payoffs than any pure strategy. Thus, when we reevaluate the decision problem during execution, we should consider all behavioural rules.

Proposition 1.6.4 *A strategy is modified multiself consistent if it is multiple multiselves consistent. A strategy is multiple multiselves consistent if it is time consistent.*

The proof of Proposition 1.6.4. The requirement of modified multiself consistency is one of the conditions in the definition of multiple multiselves consistency. Thus, a strategy is multiple multiselves consistent if it is modified multiself consistent.

By the second condition in Definition 1.6.3, if a time consistent strategy σ is not multiple multiselves consistent, there exits a behavioural strategy $\sigma' \neq \sigma$ such that $\exists h \in H$,

$$\sum_{h\in X} \mu(h|X;\sigma) \sum_{z\in Z} p(z|h,\sigma')u(z) > \sum_{h\in X} \mu(h|X;\sigma) \sum_{z\in Z} p(z|h,\sigma)u(z).$$

It contradicts to the definition of time consistency. Thus, a strategy is multiple multiselves consistent if it is time consistent. Q.E.D.

Besides modified multiself consistency and time consistency, we could also compare our multiple modified consistency with other concepts regarding consistency. The constrained time consistency in Battigalli (1997) states that the decision maker will not reconsider the decision problem unless he reaches at a node where he should not reach by implementing his initial strategy. A strategy is constrained time consistent if it achieves the highest expected payoffs among the strategies which are different from the initial strategy at the actions assigned at histories that cannot be reached by implementing the initial strategy. In the conclusion of this paper, it states that "a strategy is constrained time consistent if and only if every equivalent strategy is a modified multiself equilibrium. Thus, combined with proposition 1.6.4, a strategy is constrained time consistent if it is multiple modified consistent. The game tree time consistency in Halpern (1997) explains that a decision maker reconsiders at a unique information set during execution, and only does the reconsideration at the first node he reaches the picked information set. The decision maker could change his strategy at that node. A strategy is game tree time consistent if the decision maker has no motivation to deviate when he reaches that node. Halpern proves that if a decision problem only contains absentmindedness but no other types of imperfect recall, a strategy is ex-ante optimal if and only if it is game tree time consistent. Thus, with the same assumption of a decision problem, a strategy is multiple modified consistent if and only if it is game tree time consistent. In fact, we could compare our multiple multiselves consistency with other concepts of consistency through its relation with ex-ante optimality and modified multiself consistency.

Gilboa (1997) also proposes the multiself approach. In his model, nature assigns equal probability to the situation that the current mover actually moves firstly in the decision problem tree and the situation that the current mover actually moves secondly in the decision problem tree. In that way, he interprets the decision problem with absentmindedness as a two-agent game with perfect recall. In fact, the equilibrium results would be changed even if we assume different probabilities to be one of those two cases. There is no planning stage in his approach. However, the backward is that it is different to apply the approach to an example in real life. Another issue is that the same as PR, Gilboa (1997) examines one-shot deviation while the PMEC in our approach indicates an ex-ante optimal strategy is not only modified multiself consistent but also stable in terms of multiple one-shot deviations.

1.6.4 “Time Consistency" Issue in Psychological Multiself Approach with a Knowledgable Planner

So far, we have dedicated a significant portion of this paper to discuss the psychological multiself game form $PG(\Gamma)$ with a confident planner of a decision problem Γ. It motivates by the fact that a strategy is ex-ante optimal if and only if it is time consistent. However, the equivalence breaks if imperfect recall presents. Then, we would like to explore to what extent of the decision maker's control power during execution makes an ex-ante optimal strategy could maintain its conditional optimality. The answer to it is that the decision maker will not deviate from the ex-ante strategy if he is allowed to have multiple one-shot deviations.

For a decision maker with perfect recall, it does not matter whether he chooses an optimal strategy ex-ante and follows it during execution or he chooses the conditionally optimal current behavioural rule when he reaches each history. The decision maker could either be a planner or the doers. However, when incorporating the stage of planning into our approach, the decision maker must be both a planner and doers, which makes the procedure of decision change from a one-step process to a two-step process. It seems to increase the complexity of the decision making. From the aspect of computational algorithm, it increases the capacity of stored information. However, for those not large decision problems with imperfect recall, we are motivated to develop the psychological multiself approach.

The main motivation is that our approach is closer to describing the process of decision in the situation of imperfect recall (absentmindedness) in real life than the classic decision theory. Imagine that you are a decision maker who faces a decision problem that involves multiple successive identical repeated situations. If you know that the future you (you in the process of decisions) will present imperfect recall (absentmindedness) and will be indistinguishable among several situations, you might choose to make a plan before the process of decisions. You know that the plan you make ex-ante would be a reminder for the future you during execution no matter the future you will directly follow the ex-ante plan or choose a conditionally optimal immediate action based on the hint of the ex-ante plan. Especially in the situation of absentmindedness, the decision maker cannot deliver either he has passed a node or his strategy at that node if the node is in an information set presenting absentmindedness. Then, for a decision maker who has limited time such that he cannot analyse how he would act if he is at other nodes in the same information set, or for agents in the manager-agent model whose communication with each other is hard to be clear and precise, it could help a lot if there is an ex-ante plan.

In classic decision theory, a rational decision maker with imperfect recall behaves the same way as if there is no imperfect recall presenting in the decision problem except for the restriction that the behaviours at nodes in the same information set should be identical. Is it reasonable if the decision maker evaluates the decision problem ex-ante in the same way no matter whether he notices the presence of future imperfect recall? Should the decision maker behave differently from that described in classic decision theory once he realises the appearance of future imperfect recall and the influence of his chosen strategy on his actions at information sets with imperfect recall? The psychological multiself approach with a knowledgable planner provides a viewpoint to answer these questions.

The “time consistency" issue is at the centre of our discussion. We could see from the example of the absent-minded driver paradox that a PMEK is not “time consistent" in a large probability. The “time consistency" here presents the way that the equilibrium strategy of the planner is the same as the equilibrium collective strategy of doers. Then, can the planner choose a suboptimal strategy but restrict the doers’ equilibrium collective strategy to be the same as this strategy?

Of course, a knowledgable planner could choose a suboptimal strategy to make the doers’ collective strategy identical to this suboptimal strategy. However, it is transferred to the same situation with a confident planner. The existence of a PMEC requires that the doers’ equilibrium collective strategy is identical to the planner’s equilibrium strategy. The behaviour belief of a knowledgable planner is the doers will play a strategy from their best responses to each other and the known planner's strategy in equilibrium. Thus, if a knowledgable planner chooses a strategy which is not optimal but "time consistent", the strategy would be one of the ex-ante strategies of the decision problem Σ. Back to the absentminded driver paradox, the knowledgeable planner's strategy which is suboptimal but time consistent is $(2/3, 2/3)$.

1.7 Conclusions and Extension

In this paper, we develop a psychological multiself approach to analyse decision problems with imperfect recall (mainly absentmindedness). The story starts from the fact that an ex-ante optimal strategy is not time consistent. It indicates an ex-ante optimal strategy is not conditionally optimal if the decision maker is allowed to reconsider once during execution and deviate to any strategy. Then, PR presents that an ex-ante optimal strategy is conditionally optimal if the decision maker is allowed to reconsider once and deviate to any one-shot behavioural rule. A strategy satisfying the above conditions is defined to be modified multiself consistent. Thus, we would like to explore whether an ex-ante optimal strategy is conditionally optimal if the decision maker is allowed to reconsider multiple times along one round play of the decision problem and deviate to a one-shot behavioural rule at each time he reconsiders.

To evaluate one-shot deviations, a multiself approach is a proper model to discuss one-shot deviations. We model each node to which the decision maker moves a distinct self and call it a doer. A doer evaluates the strategy based on the conditional expected payoff function at the node where the doer is located. If a doer is located at a node in an information set presenting imperfect recall, he should form the position belief on the probability distribution of being at every node in that information set. The nature of absentmindedness makes it impossible to deliver the nodes the decision maker passes and the strategies chosen at those nodes if the nodes are in the information set with absentmindedness. Therefore, we add an extra self into our approach, the planner. The planner moves before the decision problem starts and chooses a complete strategy for the decision problem. The planner is a passive player in the psychological multiself game form of the decision problem. He does not implement any actual moves but influences the doers' actions through their position belief. The doers form their position beliefs based on the strategy of the planner. The planner acts before all the doers and the strategy of the planner is the common knowledge of doers. The planner chooses the strategy according to his behaviour belief on the doers' behaviours. A confident planner believes the doers follow his chosen strategy. A knowledgable planner believes the doers reevaluate the decision problem and choose their conditionally optimal one-shot behavioural rule. Thus, the knowledgable planner believes a doer plays one of his best responses to other doers' strategies. With different beliefs on the doers' strategies, the planner's utility function is a psychological utility function.

We define the concept of a psychological multiself equilibrium with a confident planner (PMEC) and a psychological multiself equilibrium with a knowledgable planner (PMEK). In a PMEC, the equilibrium strategy of the confident planner, the equilibrium collective strategy of doers, and the confident planner's first order belief on the doers' strategies are the same. The PMEC illustrates that an ex-ante optimal strategy is conditionally optimal even if the decision maker is allowed to reconsider multiple times during the execution. Additionally, with the precise that the decision maker forms position belief using his chosen ex-ante optimal strategy, the ex-ante optimal strategy is stable. The stability reflects the fact that there does not exist a group of nodes where the decision maker reconsiders and deviates such that the decision maker's conditional expected payoffs at any node strictly increase. The planner in a PMEK presents the decision maker's different ways of evaluating a decision problem with imperfect recall and a decision problem with perfect recall. It is also the motivation of our approach since the decision maker in the conventional game theory treats perfect recall and imperfect recall in the way except for the restriction of identical behavioural rules at nodes in the same information set. We expect the decision maker's behaviours should be changed if he notices the appearance of future imperfect recall.

The contribution of our approach is as follows. Firstly, we apply psychological games to analyse decision problems with absentmindedness. Psychological games normally model emotions or the interaction among individuals. It does not mean psychological games can only be used to model psychological factors. Strictly speaking, psychological games are used to model belief-dependent preference. The belief-dependent preference implies an individual's preference changes as his beliefs about others' behaviours change. In our approach, the planner's preference over the complete strategies of the decision problem also changes as his behaviour belief on the doers' behaviours. Thus, it is reasonable to incorporate psychological games into our approach.

Furthermore, by different beliefs of the planner, we attempt to model the way of evaluating the decision problem ex-ante when a decision maker notices the appearance of future imperfect recall. In principle, it should be different from the way the decision problem presents perfect recall. Besides those, our paper examines to what extent an ex-ante optimal strategy is conditionally optimal. The answer to it is that the ex-ante strategy is conditionally optimal even if the decision maker reconsiders the decision problem several times during one round of processing. The last point of contribution is that the strategy at an information set presenting absent-mindedness could determine the decision maker's position belief, and at the same time, the belief influences the decision maker's choice of strategy. The cycle of "strategy-belief-strategy" is unreasonable. However, in our approach, the doers are assumed to form position belief by the planner's strategy and it separates the process of position belief forming and the strategy choosing. By introducing a planner, the approach solves the cycle problem.

However, there are restrictions on our approach. Firstly, by introducing a planner into our model, we transfer the decision problem from a one-step procedure to a two-step procedure. Although our approach might be closer to describing the situation in real life, it increases the complexity of the decision making. In the aspect of computational algorithms, more data needs to be stored in our approach. We only propose two beliefs of the planner. There are many possibilities of beliefs to be explored. Despite those restrictions, our approach is a good attempt to analyse decision problems with absentmindedness.

Chapter 2

Analysis of Decision Problems with Imperfect Recall Using IS Expected Payoff Function

2.1 Introduction

In classic game theory, people are assumed to have perfect ability in the process of decision making. For instance, when uncertainty appears, people learn not only all states they will probably meet but also the probability distribution of each state. Another example can be that people are assumed to remember their past experiences. However, most people only have limited capabilities of reasoning and remembering in the real world. A concrete illustration can be the game of poker. There are few players who are able to remember all the poker cards which have been shown up and who play those cards. Most people are able to remember solely a part of those memories. We say people have imperfect recall when describing the fact that people are far from perfection in memory.

Imperfect recall is a representative of bounded rationality. A player is defined to have **imperfect recall**, if he forgets any of his past actions or any information he has previously acquired. Normally, we could use imperfect recall to model the situation in that people are confused about several situations due to the forgetfulness of their past actions. Most papers in game theory exclude situations with imperfect recall in their discussion over solution concepts and properties. A reason why it is excluded from the mainstream analysis of games is that games with imperfect recall is much harder to interpret than those with perfect recall. The interpretation in different aspects which has no essential differences under perfect recall become significant when taking imperfect recall into account, such as single deviation (if absentmindedness presents), the outcome-equivalence of mixed strategy and behavioural strategy. However, it is necessary to model imperfect recall since most people present imperfect recall. For example, the fact that people cannot know the change in their saving physically and visually might cause their overspending. Imperfect recall could also be used to model a team composing numbers of agents who complete a task together with a lack of communication among themselves.

Besides various difficulties in interpretation of games with imperfect recall, a fundamental problem is that a player might be not able to guarantee the complete implementation of a previously chosen strategy along the execution, even without any new factors (new information regarding the game, the change of player's preference, etc.) intervene. However, in games with perfect recall, every player does not have any different profitable strategy from the previously chosen one when he reevaluates the game during play. We say a strategy which satisfies the above description is time consistent. The word "time" in this term refers to the procedure of a game rather than intertemporal issues. Games which exhibit time inconsistency problem could be totally irrelevant to intertemporal choices. According to Osborne & Rubinstein (1994), time consistency is called structural consistency in a general extensive game. Alternatively, we could also understand it as the examination of sequential rationality at reachable information sets. The potential deviations from the ex-ante chosen strategy during execution make it difficult to expect what behaviours could be observed once the game starts and what would be the results of the game in the aspect of equilibria, which is one of the important objects of game theory. Thus, we should analyse games with imperfect recall using a new framework, in which the definition and properties of solution concepts should agree with those in classic theory with the assumption of perfect recall. To simplify the problem, we focus on discussions of decision problems, one-player game, in this paper.

We find that, in conventional decision problem theory, there is not much difference between the following two situations when a decision maker reconsiders a finite decision problem with imperfect recall. The first one is that the decision maker knows he will present imperfect recall at a later information set. The second situation is that the decision maker does not know he will present imperfect recall but he is required to choose the same behavioural rule at those nodes which are in the same information set. Then, the assumption that the decision maker knows the structure of the decision problem tree is meaningless. In other words, it seems to be indifferent whether a decision maker knows he will present imperfect recall as the decision problem executes. It is interesting to explore how a decision maker would behave if he prepares for his future imperfect recall when he reconsiders during execution. We call such a decision maker sophisticated.

In a finite decision problem with perfect recall, the conditional expected payoff function could be written as a recursive form such that, for an information set X and all information sets X_j which are reached by one-shot move from a node in the information set X, the conditional expected payoffs at X is the sum of the probability of reaching X_j from X multiplies the conditional expected payoffs at X_j. The conditional expected payoff function at an information set X contains the information on how a decision maker considers the decision problem. Then, we wonder whether it is possible to develop a recursive calculation rule to model how a sophisticated decision maker reconsiders the decision problem with imperfect recall during execution.

Therefore, a sophisticated decision maker at an information set explains his current one-shot behaviour leads him to another information set instead of a history (the one-shot behaviour could also leads him to the same information set if the current information set where he is located presents absentmindedness). The attitude modified IS-path probability function in the sophisticated recursive calculation rule describes the situation that how a decision maker evaluates the decision problem at an information set X. Firstly, he forms a belief system consistent with the strategy σ he is executing now. The belief system indicates the probability of being at a node $h \in X$ conditional on the information set X. Conditional on being at the node h, he could actually implement the action a. The probability that he implements the action a is described by the strategy σ. Once the decision maker implements the action a, it could lead him to either the same information set if information set presents absentmindedness or another information set. If the action leads him to the same information set, the decision maker expects to obtain the full conditional expected payoffs. If the action leads him to another information set, the decision maker knows it is possible that he could also reach that information set through a path which does not contain any node from the information set X. Then, the decision maker expects λ of the conditional expected payoffs at that information set. The parameter λ conveys the viewpoint that the decision maker at the beginning expects to obtain identical expected payoffs of an information set no matter which path he passes through. By this step, the attitude modified IS-path probability function calculates the attitude modified probability from the information set X to each information set which the decision maker can reach with one-shot move.

If we repeat this step at those information sets which can be reached by the decision maker with one-shot move, it calculates the probabilities to reach other new information sets. If we keep repeating this step until a terminal node. By recursion, it calculates the attitude modified probability of an IS-path from the information set X to any terminal node. We say there is an IS-path from information set A to information set B if, for any adjacent two information sets in the IS-path, the decision maker could reach one of those two information sets to the other one by one-shot move. Then, there is another problem that a terminal node could be considered in more than one IS-path. However, there is at most one IS-path containing the real path from the decision maker's current information set to any terminal node. Some terminal nodes that will never be reached conditional on the current information set should also be considered currently. It is because the decision maker knows the structure of the decision problem and the fact he will lose the information of the current history he is located if the information set reached by his one-shot move presenting imperfect recall. The terminal nodes considered at those information sets he thinks he might reach in the future due to imperfect recall should also be considered now.

The average function solves the problem of multiple IS-paths from an information set to a terminal node. Since the decision maker acts as if he cannot figure out which IS-path contains the real path to a terminal node when there is more than one IS-path, the decision maker is assumed to assign equal probability to each one. Then, we apply the attitude modified IS-path probability function to the substitute function. We call the resulted conditional expected payoff function IS expected payoff function. The IS expected payoff function implies the way a sophisticated decision maker considers conditional on an information set. We are going to prove that a strategy is ex-ante optimal if and only if it is IS-time consistent for a finite decision problem without crossing information sets. Two information sets X_1 and X_2 are defined to be crossing information sets, if $\exists h_1, h'_1 \in X_1$, $\exists h_2, h'_2 \in X_2$, such that h_1 is a subhistory of h_2, and h'_2 is a subhistory of h'_1. A strategy is IS-time consistent if it is conditionally optimal at any information set if the decision maker evaluates by the IS expected payoff function (there are some requirements for an available deviation strategy at an information set, see section 6). It implies that the decision maker cannot find a more beneficial strategy than the optimal strategy he chooses before the decision problem starts if he prepares for his future imperfect recall when he evaluates the decision problem at any information set. In other words, in the middle of a decision problem, an individual never denies the best plan he chooses ex-ante if he always takes how he will evaluate conditional on each future step into consideration when he evaluates now.

The rest of this paper is organised as follows. In section 2.2, we explain the concepts related to imperfect recall and time consistency by examples. The difficulties of formulation decision problems with imperfect recall are also discussed. The two types of expected payoff function are compared in section 2.5. In that section, we formulate the sophisticated recursive calculation rule when a decision maker reconsiders at an information set and compute the IS expected payoff function according to the calculation rule. By applying both expected payoff functions to examples, we could see the different conditionally optimal strategies at the same information set in the same decision problem. In section 2.6, we discuss the concepts of IS-time consistency and one information deviation property. Cross-branch information sets are defined with respect to the special relationship among information sets in terms of imperfect recall. In sections 2.7 and 2.8, The concept of strongly IS-time consistency is defined. The strongly IS-time consistency requires a strategy is not only optimal at the optimal paths but also at the off-optimal paths. We demonstrate the deviations at multiple information sets, and thus, define the concepts of (strongly) modified modified IS-time consistency. In the last section, we do the conclusion and refer to other types of imperfect recall.

2.2 Imperfect Recall and Time Inconsistency

For a decision problem presenting imperfect recall, there might be multiple situations. We use examples to respectively illustrate three main aspects: forgetting previous actions, forgetting previous acquired information and absentmindedness. As in example 1 (see Figure 3.4), the decision maker at information set K forgets his previous action at history e_1. The decision maker in example 2 (see Figure 3.5) is initially informed of the results of chance player at histories e_1 and e_2. However, he forgets how the chance player played when he arrives at information set X. In this case, the decision maker forgets his previous acquired information.

The third example (see Figure 3.1) is the famous " absentminded driver paradox" proposed by Piccione & Rubinstein (1997a) (PR from now on). The story behind is described as follows,

A forgetful driver plans his drive on the highway in order to come home. He will arrive home if he exits the highway at the second intersection (with payoff 4). If he exits at the first intersection, he will reach an unknown place (with payoff 0). If he keeps on driving after passing the second intersection, he is not able to turn back and thus will never arrive home, but the good news is that he could find a hotel at the end of the highway (with payoff 1). It is easy to see that to exit at the second intersection is the best choice for the driver while to exit at the first intersection is the worst. Because he is forgetful, the driver knows that he will not be able to remember how many intersections he has passed by. Besides, the environment at both intersections is very similar and there is no sign on the highway to remind the driver of his current position.

This example illustrates the problem of absentmindedness in decision making. We say an information set I presents absentmindedness if there are at least two histories in I along one path in the tree structure of the decision problem. The decision maker in I could treat the situations in different histories as the identical scenario and does not know his current history. At the moment that the decision maker reaches the second intersection, he forgets the fact that he has previously made a choice. As shown in Figure 3.1, when the decision maker reaches the information set I, he is not able to figure out his current location at either history e_1 or e_2. At node e_2, the decision maker does not know not only which action he chooses at e_1 but also whether he has made a move in the information set I. In fact, absentmindedness is the most unconventional case of the three situations. It causes much more complexity in the interpretation of decision problems.

2.2.1 Time Inconsistency in Imperfect Recall (Examples)

In the second example (see Figure 3.5), the only ex-ante optimal strategy for the decision maker is to execute G at history e_0, E at history e_1, D at history e_2, R at information set X, which yields the expected payoff of 4.

However, when the decision maker reaches history e_1, he would like to change his following behavioural rule to execute D at history e_1 and L at information set X since he knows he will receive payoffs of 2 if he resists to execute E at history e_1 but payoffs of 5 if he plays his new behavioural rule. Thus, the ex-ante optimal strategy is not time consistent once the decision maker reaches history e_1.

For the example of "absentminded driver paradox" (see Figure 3.1), we define p as the probability of executing C. Then, the only ex-ante optimal behavioural strategy is $(p^\star, p^\star)$ such that

$$p^\star = \underset{p}{\operatorname{argmax}} \left(p^2 + 4p(1-p)\right).$$

It is easily to know that $p^\star = 2/3$. However, when the decision maker is at the information set I, his optimal strategy is different from the ex-ante one.

How do we calculate the conditional expected payoffs on a non-singleton information set? It is naturally to define it as a weighted sum of conditional expected payoffs at each history within that information set. Before the discussion on what should be the reasonable weight, we assume a general weight to different histories. As in the example of absentminded driver paradox, the weight assigned to history e_1 is assumed to be α and history e_2 is $1-\alpha$, then the conditional expected payoff is

$$\alpha(4p(1-p)+p^2)+(1-\alpha)(4(1-p)+p).$$

It is easily to conclude that the optimal p is $2/3$ if only if $\alpha = 1$. In other words, the decision maker's optimal strategy at the information set X is the same as ex-ante optimal one only if he knows he is currently at history e_1 for sure. However, it does not make sense for a decision maker to know his accurate position in a non-singleton information set. It is contradicted to the natural definition of a non-singleton information set in which the decision maker should be not able to differentiate different histories and do not know his current history.

In fact, the weight assigned to different history could be the relative probability of being at that history. PR imitate the consistent belief defined in games with incomplete information to calculate the relative probability on an information set presenting imperfect recall. They propose the reasonable assumption of the relative probability is equal to the long run proportion of times in which visiting the information set involves being at each history for a decision maker who plays the decision problem numbers of times following a strategy. Therefore, the reasonable weight, generated from the ex-ante optimal strategy $p^\star = 2/3$, assigned to the history e_1 in absentminded driver paradox should be

$$\alpha = \frac{1}{p^\star + 1} = \frac{1}{2/3+1} = \frac{3}{5} \neq 1.$$

Then, the decision maker's optimal strategy at information set X has little chance to be the same as the ex-ante optimal one. Actually, the ex-ante optimality is not required in the definition of time consistency. Time consistency describes a strategy which satisfies the property should be conditionally optimal if being reexamined at any information set as the play unfolds, but not necessarily optimal before the decision problem starts. Even though the example of absentminded driver shows the inconsistency in strategy selection between ex-ante and during execution, we still say it presents the time inconsistency problem. It is mainly because, in decision problems with perfect recall, a strategy is ex-ante optimal if and only if it is time consistent; while, in this example, the only ex-ante optimal strategy is not time consistent. In this case, the timing that the decision maker chooses a strategy becomes important since the chosen optimal strategy would be different between the situation that he selects before the play unfolds and the situation of being at some information set during the play. We find that the expected payoffs at the initial node in the decision problem are the same as the ex-ante expected payoffs in decision problems with perfect recall since the decision maker at the initial node could realise any outcome through his following actions. On the contrary, the initial node is included in a non-singleton information set in this example. Once there is a positive probability of currently being at node e_2, the decision maker losses the possibility to exit at the first intersection. Therefore, the equivalence of optimality between ex-ante situation and execution is invalid in decision problems with imperfect recall. The chosen strategy might be different when an individual makes decisions at different stages of a decision problem, which causes the different expectations of the game and the different results from the execution of the decision problem.

2.2.2 Time Inconsistency in Imperfect Recall (General Explanation)

From examples 2 (see Figure 3.5) and 3 (see Figure 3.1), we know that time inconsistency could occur when the decision maker forgets previous acquired information or presents absentmindedness. How do we differentiate the three aspects of imperfect recall? The feature of absentmindedness is the most distinct of the three. An information set presents **absentmindedness** if there is a path along which cut the information set more than once. For the other two situations, the crucial point is that the player who acts at the longest common subhistory of all the histories in a non-singleton information set. If it is the chance player (or another player different from the decision maker) who acts at the longest common subhistory, the decision problem presents the situation that the decision maker forgets his previously acquired information. Conversely, if it is the decision maker who moves at the longest common subhistory, the game tree indicates that the decision maker does remember he has made actions but forgets what those actions are. The differences among the three situations are easier to distinguish if a decision maker is allowed to record his past experiences. In example 1, the decision maker could record the past strategy he chooses and the corresponding node where a strategy is proposed. In example 2, the decision maker writes down the past strategy but not necessarily the corresponding node. However, in example 3, the decision maker is not able to record either the past strategies or the nodes. Notice that, the situation where the decision maker is not informed of the actions of the chance player belongs to imperfect information but not imperfect recall. It is hard to distinguish the situation of imperfect information and that of forgetting previously acquired information (imperfect recall) only by recognising which player moves at the longest common subhistory. At a non-singleton information set X, the decision maker has imperfect information rather than imperfect recall if the information sets appear in the experience of each history in X are the same and the frequency of each information set in the experience is also the same. Otherwise, the decision maker has imperfect recall.

The elements in the set of terminal nodes that can be potentially reached conditional at an information set become strictly fewer along the play of a decision problem. With the presence of perfect recall, the set of terminal nodes which can be reached by moves from the current information set is the same as the set of terminal nodes which the decision maker should consider at the current information set. Outcomes are gradually excluded from the decision maker's consideration set as he moves. However, in decision problems with imperfect recall, it is different. The number of terminal nodes which the decision maker needs to consider might not be reduced as the decision problem unfolds but reversely increased although the set of terminal nodes could be reached is reduced. As in example 2, we only consider the behaviour at e_3 when calculating the expected payoffs at history e_1 while we evaluate the behaviours at both e_3 and e_4 at the same time when calculating the expected payoffs at information set X. In the meanwhile, the behavioural rule at histories e_3 and e_4 should be identical and the conditional expected payoffs at histories e_3 and e_4 are the same. Thus, the problem to choose the behavioural rule at history e_3 is different when the decision maker is at history e_1 and information set X. However, the decision maker knows the fact that he cannot differentiate histories e_3 and e_4 when reaching information set X prior to the play of the decision problem. We might ask why the decision maker does not take the fact of loss of memory at X into consideration when choosing an optimal strategy at e_1.

Therefore, the decision maker at an earlier stage reevaluates the decision problem in the same way no matter whether he has perfect recall or imperfect recall in the future stage. The extra information that he will present imperfect recall later in the decision problem has no effect on his current decision. At a singleton information set, i.e., a history h, the decision maker restricts his attention to the histories that might be reached from his current history. Assume one (and only one) of those histories is included in an information set X with imperfect recall, denoted by h'. The decision maker at h only considers $h' \in X$ but ignores other histories in X, even though he knows he will be confused if he reaches X in the future and considers all the histories in X at that moment. Then, it might arise our interest that what would happen if the decision maker takes advantage of the information that he will present imperfect recall later when reexamining the decision problem. Then, it seems reasonable if we regard an information set as a whole unit when the decision maker is not in that information set. If the decision maker is currently in that information set, he forms a belief system, the subjective probability of being at each history, and makes decisions according to the belief system.

2.2.3 Difficulties in Interpretation with the Presence of Imperfect Recall

In decision problems with perfect recall, the decision maker remembers his previous actions and acquired information regarding the play. A strategy is ex-ante optimal if and only if it is time consistent. The decision maker remembers his initial strategy and the latest strategy when reconsidering the decision problem during the execution. Even if we relax the assumption that the decision maker can remember his initial strategy, it is practicable to reconstruct the ex-ante situation and compute the optimal strategy, and then follow the deduced ex-ante optimal strategy for the rest of the decision problem. If the ex-ante optimal strategies are not unique, moves from any choice of those optimal strategies will satisfy time consistency.

One-shot Deviation Principle

In a finite decision problem with perfect recall, a strategy is time consistent if and only if it satisfies one-shot deviation principle. One-shot deviation principle demonstrates the decision maker cannot increase his expected payoffs by deviating to a different behavioural rule of the current information set from his previously chosen strategy when he is asked to reevaluate during execution, while assuming the behaviours at other information sets are immutable. In a finite decision problem with perfect recall, a strategy is time consistent if and only if it is ex-ante optimal. Thus, in a finite decision problem with perfect recall, a strategy is ex-ante optimal if and only if it satisfies one-shot deviation principle. At every information set, the decision maker has no better choice than following the chosen ex-ante optimal strategy no matter how many moves he can control for the rest of play. Furthermore, how often the decision maker thinks about deviations and whether the decision maker knows where he does the reevaluation in the game tree are not crucial questions in the environment of perfect recall.

However, if imperfect recall, especially absentmindedness, appears, most of the insignificant questions becomes critical. Firstly, which is also the motivation of our paper, an ex-ante optimal strategy is not time consistent. Furthermore, a strategy which satisfies one-shot deviation principle might not be time consistent. For example, see Figure 3.6, the strategy $s = (e_0, e_1, e_2, e_3) = (T, R, r, r)$ satisfies no profitable deviation at any information set but is not time consistent since the behavioural rule $(e_1, e_2, e_3) = (L, l, l)$ can realise a higher payoff than the strategy s when reexamining at the node e_1. However, a strategy must satisfy one-shot deviation principle if it is time consistent.

A Decision Maker's Control Power

The control power of a decision maker matters with the presence of imperfect recall since time consistency is not equivalent to no one-shot deviation in such a scenario. The concept of modified multiself consistency in PR shows an ex-ante optimal strategy satisfies the principle of sequential rationality at every information set which can be reached with positive probability by implementing the strategy if the decision maker's control power is restricted to his immediate move. A crucial implication of one-shot deviation principle could be that the decision maker cannot update his current strategy but remember the correct initial strategy he chooses before the play. The decision maker is aware that he will not notice the switch to a new strategy in the later stage during execution even if the conditional expected payoff of the switched strategy is higher than the initial one. It is meaningless to consider a strategy for the rest of the decision problem if the updated strategy cannot be conscious if the decision maker moves to a new information set. In that case, one-shot deviation should be the most proper model to analyse decision problems with imperfect recall.

Multiple Optimal Strategies

When we calculate the expected payoffs at a non-singleton information set, it is necessary to form a subjective belief system, i.e., the probability distribution of being at each node in that information set, which is consistent with the strategy that the decision maker currently uses. Time consistency defined in PR's paper requires the decision maker remembers his initial chosen strategy and forms his belief according to it. The same as the situation of perfect recall, the decision maker can deduce the ex-ante optimal strategy even if he forgets it at some information set. However, it becomes much more complicated if there is more than one ex-ante optimal strategy. Different behavioural rules at previous nodes might result in different belief systems at a non-singleton information set. It could be very hard for a decision maker who forgets his initial strategy to form a belief system which is consistent with it since it is possible he forms a belief system based on another ex-ante optimal strategy.

Reconsideration Related Issues

Besides the above discussions, several questions related to reconsideration play an important role in the interpretation of decision problems with imperfect recall. Those questions include but are not restricted to the following. 1) How often is the decision maker allowed to reconsider? Once in each play of the decision problem or at each reachable information set, for example. 2) Where is the decision maker allowed to reconsider? At any reachable information set or at information set with perfect recall, for example. 3) Who decides where the decision maker can reconsider? An external process or the decision maker himself, for example. 4) Does the decision maker know where he reconsiders before he chooses a strategy before the decision problem starts? Or the decision maker has the chance to do the reevaluation when he is stopped in the middle of execution?

Switch of a Strategy

Although we have demonstrated numbers of difficulties in the explanation of decision problems with imperfect recall, there are some issues that have not been discussed much. Most of the works assume the decision maker knows his current strategy during execution. However, how the decision maker notices the switch of a strategy? Battigalli (1997) proposes the concept of constrained time consistency. He explains that the decision maker notices the switch of initial strategy only if he reaches the information set which cannot be reached by the initial strategy. Halpern & Pass (2016) offer that the decision maker is able to notice the change of strategy at the previous information set only if all the histories in his current information set have a subhistory in that previous information set. The words "previous" indicates the decision problem passes the information set in the earlier stage of the decision problem. As shown in example 3, the optimal strategy at e_1 is to implement D immediately and then L instead of R at information set X, while the optimal strategy at e_2 is the same as the ex-ante optimal strategy. If the decision maker is able to notice the updated switched strategy at information set X, he must know which node he is at currently, and thus the behaviour of the chance player. The different switched strategy provides the information.

When the decision maker cannot be updated with the latest switched strategy or the decision maker is asked to reconsider his strategy more than once during one round play of the decision problem, the decision maker should be assumed to change his behavioural rule of the current information set. If the decision maker cannot remember the latest strategy, he knows he cannot follow the new strategy later even if there is a better different one at the moment of reconsideration. Thus, the decision maker can at most change his immediate behaviour. If the decision maker is assumed to have multiple reconsideration times and remembers the latest switched strategy, the updated strategy should suggest nothing but the strategy itself. For example, in example 2, even though the decision maker knows his latest strategy is D at e_1 and L at X when he is at information set X, he does not know he was at e_1. It might be caused by a potential mistake in reconsideration or the decision maker did not choose the conditional optimal strategy but a different strategy for some reason when he reconsidered at X. There are two possibilities. One is that it is the external process which picks the node where the decision maker reconsiders and the decision maker does not know how many times and where he should do the reconsideration. Then, he is not sure whether he could commit to the new strategy since he might reconsider at the next move and choose another strategy. It is naive to expect a commitment to a strategy if the decision maker is aware of a possible revision of the plan is possible. However, if the decision maker knows the nodes where he will reconsider, no matter nodes are chosen by an external process or the decision maker himself, his control power should be more than just one step. On the contrary, if the time consistency relevant concepts are discussed, the decision maker is assumed to know the switched strategy and some external process chooses a unique node where he reconsiders, and he will not reconsiders anywhere else.

It seems unreasonable to assume people are able to remember their latest strategy but present imperfect recall. Like in example 1, if the decision maker at information set K remembers the strategy he switched at node e_1 which indicates he played L at e_1, then he must deduce he is at node e_2 currently. However, when the strategy is a mixed strategy or behavioural strategy, it is not contradicted between updated strategy and imperfect recall since the decision maker might remember the behavioural rule but forget the resulting implementation of the strategy. Still in example 1, if the decision maker's latest strategy specifies he should play L and R with equal probability at node e_1. When he moves to information set K, he could remember the behavioural rule $(L,R) = (1/2,1/2)$ at e_1 but forget he implemented L or R actually. It is related to the explanation of randomisation in strategy. Although mixed strategy and behavioural strategy specify probability measure over available actions, there is only one definite action implemented during each execution.

Absentmindedness Makes it More Complicated

So far, the difficulties in interpretation of imperfect recall which have been discussed exclude the situation of absentmindedness. When absentmindedness presents, along a path, an information set and a node are not equivalent. Gilboa (1997) offers another formation of the absentminded driver paradox and transforms it into a two-agent game with perfect recall. Thus, the consideration at an information set and at a node is equivalent under his formation of absentmindedness. Lambert et al. (2019) are inspired by Gilboa and incorporate his multiself model into the analysis of games with imperfect recall.

Thus, the definition of agent form of a decision problem with absentmindedness should specify whether a node or an information set should be a distinct agent. The situation that two or more nodes belonging to a common information set along a path brings more problems than the definition of agent form. Back to the questions about reconsideration. Denote the information set presenting absentmindedness X and two nodes h_1, h_2 in the information set X belong to the same path of the decision problem and h_2 is reached by the implementation of some action at h_1. Assume that the decision maker is allowed to reconsider the play at each node in the information set X. When he is at node h_1 and considers what should be currently the optimal strategy conditional on the information set. According to the definition of information set, the decision maker does not know he is currently at h_1 or h_2 but only knows he is in the information set X. In one case, if the decision maker knows he will reconsider at h_2 again based on his belief that he is currently at h_1, he should realise he would not be able to notice the switch even if he chooses a new strategy for the whole play. Thus, his control power must be restricted to the current move instead of a strategy. Additionally, he should choose an optimal move while assuming an arbitrary behavioural rule at h_2 since there might be also a deviation at h_2 when he moves to that node. Compared, if the decision maker thinks he reconsiders only once at an information set, as assumed in the concept of time consistency and modified multiself consistency in PR, he will choose an optimal behaviour with the belief that he will follow the initial strategy if he moves to h_2. Thus, within an information set with absentmindedness, the decision maker's control power must be limited to his current move when he reconsiders. The discussion here to some extent coincides with the framework in Hillas & Kvasov (2020a). They defend that, in an extensive game form, the player's belief should include not only the probability distribution on an information set but also the strategies the others are playing. if the decision maker in node h_1 regards the future himself in node h_2 as another play, then the multiself model presents a similar analysis to Hillas & Kvasov (2020a). Furthermore, switching to a strategy becomes problematic if doing the reexamination in an information set with absentmindedness. A strategy requires the same behavioural rule at every node in the same information set. Assume that a decision maker is allowed to reconsider once in the information set X at h_2 rather than h_1. It seems to be not correct to say the decision maker has a better conditional strategy if he reconsiders at h_2 and finds out a better (assigned with higher expected payoffs) behavioural rule for the rest of the decision problem.

We have analysed why the decision maker's control power over his behaviours in an information set with absentmindedness allows at most one step deviation. It might be argued that in the example of "Forget previous acquired information" (see Figure 3.5), the decision maker at the node e_1 and e_2 should be able to change at most his current move since he forgets his previous actions, let alone a switched strategy. In fact, the decision maker can calculate the optimal strategy at h_1 and h_2 when he is at the information set X. It is problematic if the conditional optimal strategies at h_1 and h_2 are different since the decision maker does not know which node he has passed. If the conditional optimal strategies at h_1 and h_2 are the same but different from the initial one, the decision maker at the information set X could be aware of the switch. Thus, it is meaningful to consider a behaviour plan for the rest of the decision problem when the decision maker is at h_1 or h_2. However, with absentmindedness, the decision maker is not even sure whether he has reconsidered at the previous node or not if multiple times of reconsideration are allowed at an information set. It is reasonable that the decision maker can only control his current behaviour when he reconsiders in an information set with absentmindedness once he is aware of this fact.

Mixed Strategy or Behavioural Strategy?

Kuhn (1953) proposes that, in a finite game with perfect recall, there is a behavioural strategy which is outcome-equivalent to any mixed strategy, and vice versa. However, in the presence of imperfect recall, the outcome-equivalence between mixed and behavioural strategies is not valid anymore. For example, in the example of absentminded driver paradox, the two pure strategies yield the payoffs of 0 (E at both intersections) and 1 (C at both intersections) while the ex-ante optimal behavioural strategy $p^\star = 2/3$ yields a higher expected payoff $4/3$. PR have proved that there exists a behavioural strategy that yields strictly higher expected payoffs than any pure strategies (mixed strategies). Thus, we focus on behavioural strategies in the paper since it is less relevant to the existence of equilibrium in the discussion of decision problems.

Since time consistency examines the optimality of a strategy when the decision problem proceeds, we should analyse mixed strategy and behavioural strategy from the perspective of execution. The mixed strategy is the randomisation over different pure strategies, and the randomisation occurs before the decision problem starts. The decision maker's behaviours should be definite during the play if the decision maker chooses to implement a mixed strategy. Compared to behavioural strategies, mixed strategies indicate decision maker's coordinated behaviours across information sets. It might arise a question that, if the decision maker is assumed to know the latest strategy, does he know the realised pure strategy from the support of the mixed strategy he chooses before the decision problem begins? Notice that the support of a mixed strategy is the set of pure strategies on which the probability measure is positive. In the definition of time consistency, we compare the expected payoff of the current strategy with that of different pure strategies. The reason why we do that is not only the expected payoffs of a mixed strategy cannot exceed the payoffs of each pure strategy from its support but also the decision maker can only deviate to a pure strategy during execution.

With the presence of absentmindedness, the definition of time consistency should be revised. As we have stated, there might be a behavioural strategy which assigns higher expected payoffs than any pure strategy. Thus, in the definition of time consistency, we should compare the “baseline" behavioural strategy with any behavioural strategy instead of a pure strategy. A behavioural strategy assigns independent randomisation at each information set, it realises much less coordination of actions across different information sets.

Multself Approach and Agent Equilibrium

The agent form of an extensive finite decision problem splits every player in the game into multiple agents, one for each of his information sets, and all agents of a player are assigned the same payoffs as the player (Kuhn 1953). A strategy is an agent equilibrium if it is a Nash equilibrium of the agent form of the decision problem. Every strategy which satisfies one-shot deviation principle must be a Nash equilibrium in the agent form of a decision problem. No matter whether the decision maker chooses his immediate behaviour before the play or during the execution, the same optimal behavioural rule will be chosen in decision problems with perfect recall.

Strotz (1955) proposes the multiself approach. A strategy is a multiself equilibrium (mutiself consistent) if the decision maker at each history is modelled as a different self, every self of the decision maker cannot increase his expected payoffs by changing his immediate behavioural rule when he is called upon to move. In a decision problem with perfect recall, a strategy satisfies one-shot deviation principle if and if it is multiself consistent since a path can at most cut each information set once. During execution, it is indifferent to the decision maker to change his behavioural rule at the current information set or history.

However, when absentmindedness appears, there might be more than one history in the same information set along a path. Thus, the multiself approach and one-shot deviation principle describe different scenarios at an information set with absentmindedness. An equilibrium in multiself approach indicates that the decision maker has no profitable deviation at the current move while following the previously chosen strategy at other histories in the same information set. On the contrary, the one-shot deviation principle assumes the decision maker can control his behaviours at any history in that information set. The two approaches describe the two opposite assumptions upon reaching an information set (Piccione & Rubinstein 1997a). One is that a decision maker takes his actions to be immutable at future occurrences of that information set, no matter what actions he is implementing now. It is the idea behind modified multiself consistency. Another extreme assumption indicates that the decision maker in the information set considers changing his behavioural rule and expects that he would adopt whichever action he uses now if the same information set occurred again. The control power endowed to the decision maker is different in these two frameworks, immediate action and actions at all occurrences of an information set, respectively. We agree with Aumann et al. (1997a) that, it is different between the player expects the same action at another history and the player decides the same action at different histories.

2.3 Literature Review

Imperfect recall reveals the bounded rationality of human beings. As shown in the example of the absentminded driver paradox, the fact that ex-ante optimal strategies of the decision maker may not be time consistent makes us focus on the procedural aspects of decision making. Some progress has been made to help understand the complexity brought by imperfect recall although, compared to the exhaustive literature on perfect recall, much fewer works discuss what should be a proper framework to interpret imperfect recall.

Kuhn is the first one who distinguishes explicitly between games with perfect recall and imperfect recall. Kuhn (1950) indicates that in games with perfect recall, players remember everything they knew and did in the past at each occasion when they move. In 1953, he gives an alternative definition of perfect recall which associates this concept with strategies. Notice that, in Kuhn's definition of extensive form games, any play of a game should cut each of the player's information sets at most once, i.e., absentmindedness is excluded. Later, Isbell discards the restriction that one information set can contain at most one node from a path in a game tree (Isbell 1957, 1959). He defines extensive games which satisfy the above condition as linear games, while those which do not meet this requirement as nonlinear games, such as absentmindedness driver paradox and repetitive games (Alpern 1988).

Then, PR present the example of the absentminded driver paradox which attracts many researchers' attention. Quite a lot of them explain the example using different solutions as their replies to the ambiguities in the interpretation of imperfect recall proposed by PR (Aumann et al. 1997a, Battigalli 1997, Gilboa 1997, Grove & Halpern 1997, Halpern 1997, Lipman 1997, Piccione & Rubinstein 1997b, Segal 2000). Furthermore, they elucidate the concept of modified multiself consistency, which explains ex-ante optimality of a strategy in a multiself approach. One subtlety could be noticed that, in PR, the one shot deviation considered in modified multiself consistency is restricted to a pure action instead of a probability distribution over the available action set. It is actually far from enough to evaluate only pure behavioural rules at an information set in a general decision problem. A behavioural rule might attain higher expected payoffs than any pure action at an information set presenting imperfect recall. As Lipman (1997) points out in his paper, Gilboa (1997) and Aumann et al. (1997a) defend PR's multiself method to be as a good attempt to explain imperfect recall. However, they formulate it in slightly different ways. Aumann et al. (1997a) demonstrate that the decision maker's control power is restricted to his immediate move and he should have the same reasoning at both intersections. Additionally, the decision maker uses the behaviour at other intersection to form his belief on the current information set. They define a strategy which is optimal under such conditions satisfies action optimality, while an ex-ante optimal strategy is planning optimal. Compared, Gilboa (1997) provides an alternative formation of the absentminded driver paradox which makes it a two-identical-agent game with perfect recall. Neither agent knows if he moves first or second and both of them know they have the equal probability to be the first or second mover. Then, Lambert et al. (2019) extend the formation of Gilboa to define multiself agent equilibrium.

However, there are also situations in which one may want to give the decision maker more control over future actions. Some of the researchers argue that time inconsistency arises from the not well-defined calculation of conditional expected payoffs with the existence of absentmindedness. Grove & Halpern (1997) point out that the example of absentminded driver presents both expectation paradox and strategy-change paradox. The expectation paradox is actually asking the question of why receiving a utility *u* with probability 1 is not worth exact *u* ex-ante. They explain that it is because we do not distinguish the event of reaching a node and being at a node. Furthermore, Grove & Halpern (1997) define the calculation of expected payoffs conditional on reaching an information set and being at an information set respectively. They argue that it is more reasonable to use the subjective belief over nodes in the frontier of an information set to calculate expected payoffs conditional on reaching an information set. Halpern (1997) compares four different concepts of time consistency. One of those is game tree time consistency which calculates the conditional expected payoffs using the same formula as defined in Grove & Halpern (1997). He proposes that ex-ante optimality is equivalent to game tree time consistency. Additionally, Halpern & Pass (2016) adopt this calculation rule and extend it to a finite game with imperfect recall and define sequential equilibrium in such an occasion.

Compared with the calculation formula using the beliefs over nodes from the frontier of an information set, the two papers by Dimitri(1999, 2009) present another method to calculate conditional expected payoffs. He elucidates the assumptions of welfare symmetry and belief consistency. Welfare symmetry expounds the decision maker should examine the expected game value with respect to the whole information at some decision node. The expected game value should be the same at all histories in the same information set. Thus, it solves the "paradoxical" facts caused by absentmindedness. Our framework adopts the same technique to explain absentmindedness. However, our approach is different from Dimitri(1999, 2009) in two aspects. Dimitri excludes the situations in which the chance player works while we take it into consideration. When incorporating the expected payoffs yielding at a history, not in the information set that the decision maker reexamines, He advocates the expected payoffs of the subgame starting from that history. We propose a structure which imitates the subgame defined in classic game theory but starts with a non-singleton information set.

2.4 Necessary Features of the Definition

In the abstract structure of a finite extensive decision problem Γ, there are many nodes. A set of actions (an action set $A(e)$, an action $a \in A(e)$) is assigned to each node $e \in E$ (E is the set of those finite nodes). The decision maker chooses an action or a probability distribution over the available actions at each node. Only one action is realised even if the decision maker chooses a probability distribution over several actions.

Definition 2.4.1 *A strictly partial order $\prec$ among those nodes such that, for any two nodes e_1 and e_2, we say e_1 **strictly precedes** e_2, i.e. $e_1 \prec e_2$, if the decision maker reaches node e_2 by several moves at e_1. Denote $P(\cdot) : E \to E \cup \emptyset$ the **immediate predecessor function**, the decision maker reaches the node e by one step move from the node $P(e)$. $P_{k+1}(e) = P(P_k(e))$ the $k+1$th immediate predecessor of the node e.*

There exists a node which precedes every other node in the decision problem, it represents the start of play, denote $\emptyset$. There are also several nodes such that no nodes are preceded by them. We call a node which satisfies the above condition a **terminal node**. The set of terminal nodes is denoted by Z. There is no node x such that $P(x) = z$, for any $z \in Z$. Thus, the abstract structure looks like a tree, in which the node representing the start works as the root, the partial order among nodes compose the branches while the terminal nodes are the leaves. Alternatively, a **history** h is assigned to each node. It is interpreted as a feasible sequence of actions taken by the decision maker or the chance player. Then, the history assigned to the initial node is $\emptyset$. The set of histories is denoted by H, $\emptyset \in H$. Each history is finite since the decision problem we discuss in this paper is finite. We call h_1 is a **subhistory** of h_2 if their corresponding nodes e_1 and e_2 satisfy e_1 is a predecessor of e_2. Define C is a subset of H, and the chance player moves at the nodes in C. In this paper, we do not particularly differentiate between a history and a node. The decision maker's belief about the chance player's behaviour is denoted by ρ. ρ assigns to each history $h \in C$ a positive probability on $A(h)$. Specially, denote c the behaviour of the chance player, and $c_h \in A(h)$ if $h \in C$.

The set of information sets ($\mathscr{X}$, an **information set** $X \in \mathscr{X}$) is defined as a partition of nodes in $H - C - Z$. At a non-singleton information set, the decision maker cannot distinguish among different nodes within the set, i.e. he does not know which node he is currently at when we stop the decision problem. Additionally, the action sets assigned to the nodes in the same information set are identical, i.e. $A(e_1) = A(e_2)$, if $e_1, e_2 \in X$. There are two ways to extend the partial order on nodes to a partial order on information sets.

Definition 2.4.2 *Given two information sets X_1 and X_2, define $X_1 \prec X_2$ if for any $e_2 \in X_2$, there exists $e_1 \in X_1$ such that $e_1 \prec e_2$.*

It is easy to know that $\prec$ is a partial order in any decision problem. Halpern & Pass (2016) analyse games with imperfect recall based on this construction of partial order on information sets.

Definition 2.4.3 *Given two information sets X_1 and X_2, define $X_1 \prec' X_2$ if there exist a node $x_2 \in X_2$ which has a predecessor $x_1 \in X_1$ and $X_1 \neq X_2$.*

It is easy to see that $\prec$ is identical to $\prec'$ in decision problems with perfect recall. However, $\prec'$ might not be a partial order in decision problems with imperfect recall. For example, like Gilboa (1997) formation of absentminded driver paradox (as shown in Figure 3.7), there is no partial order $\prec$ between information set X_1 and X_2 while we have both $X_1 \prec' X_2$ and $X_2 \prec' X_1$.

Notation 2.4.1 *Denote X_h the information set containing history h.*

In this paper, we propose a framework to describe this extended order on information sets in a general decision problem.

A **payoff function** $u : Z \to \Re$ assigns a number to each terminal node which represents the payoff that the decision maker can receive when he reaches that node. The payoff function describes the decision maker's preference over different outcomes. The preference of the decision maker is assumed to satisfy von Neumann-Morgenstern axioms.

Define the **experience** of the decision maker at a history h (corresponding node e), denoted by $exp(h)$, is the sequence of information sets and actions that the decision maker experiences from the start to the the history h. We use the convention that the last element in this sequence is the information set containing history h. Then, the first element and last element in the sequence are both information sets. For example, as shown in Figure 3.7, $exp(e_1) = (X_1)$ and $exp(e_4) = (X_2, C, X_1)$. Thus, we are able to officially define the concepts of perfect recall and absentmindedness. A decision maker is assumed to have **perfect recall**, if for any two histories h_1 and h_2 in a non-singleton information set, we have $exp(h_1) = exp(h_2)$, otherwise, the decision maker has imperfect recall. It is obviously that the information set X_1 in Figure 3.7 presents imperfect recall since $exp(e_1) \neq exp(e_4)$. A decision problem exhibits **absentmindedness** if there exists a history such that the same information sets appears more than once in the experience of that history.

A **pure strategy** $s \in S$ is a function which assigns an element of $A(h)$ to each history h in the decision problem with the restriction that if h_1 and h_2 are in the same information set, the action assigned to these two histories must be identical, i.e. $s(h_1) = s(h_2)$. It works like a behavioural instruction that indicates what the decision maker should do at each decision point. The behaviour at each history is deterministic. A **mixed strategy** is a probability distribution over pure strategies. The randomisation occurs before the execution of the decision problem. Once the play starts, the decision maker implements a pure strategy, which point out a deterministic action at each information set, from the support of the mixed strategy he chosen prior to the start of play. The mixed strategy shows the decision maker's coordination ability among information sets. A **behavioural strategy** $\sigma \in \Sigma$ is a function which assigns a probability distribution over elements in $A(h)$ to each history h. Similar to pure strategies, there is restriction imposed on behavioural strategy that if h_1 and h_2 are in the same information set, the probability distribution assigned to these two histories must be identical, i.e. $\sigma(h_1) = \sigma(h_2)$. It points out the randomisation device adopted at each history. The randomisation occurs at each decision point, and it is different from a mixed strategy. As we have stated in previous section that, in games with perfect recall, a behavioural strategy has an outcome-equivalent mixed strategy, and vice versa. In other words, it does not matter that the decision maker does the randomisation ex-ante or during the execution. For a behavioural strategy, the decision maker is able to independently determine the probability distribution over available actions at each information set. It provides the foundation to establish the agent form of a decision problem. Two strategies σ_1, σ_2 are said to be **equivalent**, denoted $\sigma_1 \cong \sigma_2$, if $\forall X$ such that $p(X|\emptyset, \sigma_1) > 0$, we have $p(X|\emptyset, \sigma_2) > 0$, $\sigma_1(X) = \sigma_2(X)$, and vice versa.

Notice that we do not discuss mixed strategies in this paper. There might be ex-ante optimal mixed strategies. The payoffs assigned to the pure strategies from the support of an ex-ante optimal strategy must be the same as the expected payoffs assigned to that ex-ante mixed strategy. A mixed strategy is time consistent if every pure strategy from its support is time consistent. Thus, it is enough to explore the time consistency of behavioural strategies.

At a non-singleton information set, the decision maker cannot distinguish different nodes by observing the surrounding environment and does not know which node he is currently at. Following the calculation rule of the expected payoff at a non-singleton information set, we need to learn the probability of being at each node and the expected payoffs at those nodes. In decision problems with perfect recall, the probability of being at each node can be learnt from the randomisation device of the chance player. However, at an information set presenting imperfect recall, it might not be enough to know the randomisation device of the chance player, it is also related to the decision maker's own strategy. Thus, the decision maker should form a subjective probability distribution on the nodes by the behavioural rule of himself and the chance player. Define a **belief system** a function μ which assigns a probability to each history h in an information set X, denoted by $\mu(h|X)$, such that $\sum_{h\in X}\mu(h|X)=1$. $\mu(h|X)$ is interpreted as the probability of being at history h conditional on being at the information set X. Denote $p(h_2|h_1,\sigma)$ the probability of reaching history h_2 conditional on executing the behavioural strategy σ at history h_1. Let $p(h|\sigma)$ represents $p(h|\emptyset,\sigma)$. Then, we define a subjective belief system μ which is consistent with the behavioural strategy σ adopted by the decision maker. If the information set X is reached with positive probability by executing the behavioural strategy σ, $\mu(h|X,\sigma)$ can be interpreted as the long run proportion of times that the decision maker is at history h if he is randomly stopped at the information set X. A belief system is defined to be **weakly consistent** with the behavioural strategy σ if for every information set X which is reached with positive probability and for every $h\in X$,

$$\mu(h|X,\sigma)=\frac{p(h|\sigma)}{\sum_{h'\in X}p(h'|X,\sigma)}.$$

Furthermore, denote $p(X|\emptyset,\sigma)$ the probability of reaching information set X by executing the behavioural strategy σ. It should be the sum of probability of reaching each history in the frontier of the information set X. Denote $\hat{X}$ the **upper frontier** of the information set X, in which a history h satisfies no other history belonging to X is a subhistory of h. Thus, the probability of reaching the information set X is

$$p(X|\emptyset,\sigma) = \sum_{h\in\hat{X}} p(h|\sigma).$$

Generally speaking, $p(X|\emptyset,\sigma) > 0$ if and only if the strategy σ does not prevent information set X from being reached. We say that an information set X is relevant for σ, or **σ-relevant**, if $p(X|\emptyset,\sigma) > 0$.

2.5 Expected Payoffs at an Information Set

For a general decision problem, the decision maker's expected payoff function before the play starts (ex-ante expected payoff function) is

$$Eu(\sigma|\emptyset) = \sum_{z\in Z} p(z|\emptyset,\sigma)u(z).$$

With the assumption of perfect recall, we calculate the expected payoffs at an information set by the weighed sum of the expected payoffs at each nodes in that information set, and the weight is the decision maker's belief system. The belief system composes of probability of being at each node conditional in an information set. Denote $Eu(\sigma|X)$ the expected payoffs at the information set X, given the strategy σ; $Eu(\sigma|h)$ the expected payoffs at the history h, given the strategy σ. Then,

$$Eu(\sigma|h) = \sum_{z\in Z} p(z|h,\sigma)u(z),$$

$$Eu(\sigma|X) = \sum_{h\in X} \mu(h|X,\sigma)Eu(\sigma|h).$$

The expected payoffs at a history h is the weighted sum of the payoffs assigned to every terminal node, in which the weight is the probability of reaching each terminal node conditional at history h.

Notation 2.5.1 *Denote $Z(X)$ $(Z(h))$ the set of terminal nodes which can be reached by a completely behavioural strategy conditional on being at information set X (history h).*

A behavioural strategy is **completely mixed** if it assigns a strictly positive probability to every action available at each information set. Thus,

$$Eu(\sigma|h) = \sum_{z \in Z(h)} p(z|h,\sigma)u(z).$$

The partial order $\prec$ on information sets in decision problems with perfect recall indicates $Z(X') \subseteq Z(X)$ if for information set X, X', we have $X \prec X'$.

Definition 2.5.1 *Define $P_{IS}(\cdot)$ [1] : $\mathscr{X} \cup Z \to \mathscr{X}$ the* ***immediate predecessor information set function****. Then, $P_{IS}(X) \prec X$ and $\forall h \in X, \exists h' \in P_{IS}(X)$ such that $P(h) = h'$. If $X = P_{IS}(\{z\})$, $\exists h \in X$ such that $P(z) = h$.*

In other words, the nodes in information set X are reached by one move of the decision maker and moves of the chance player from some nodes in $P_{IS}(X)$.

The relation $Z(X') \subseteq Z(X)$ describes that, when implementing a completely behavioural strategy, the terminal nodes can be reached if the decision maker moves from the information set X must be reachable if the decision maker moves from the information set $P_{IS}(X)$. Additionally, as we have stated above, the belief system on a non-singleton information set is not determined by the decision maker. Therefore, time consistency is not a surprising property of an ex-ante optimal strategy. However, the partial order $\prec$ does not exist with the presence of imperfect recall. Instead, we develop our analysis based on the order $\prec'$. One of the differences between the relation $\prec'$ and $\prec$ is that we do not have $Z(X') \subseteq Z(X)$ if $X \prec' X'$. For example, in "forget previous acquired information" (see Figure 3.5), $Z(\{e_1\}) = \{z_1, z_3, z_4\}$, $Z(X) = \{z_3, z_4, z_5, z_6\}$ while $\{e_1\} \prec' X$. It is easy to know that the decision maker should take the outcomes z_5 and z_6 into consideration when he is at information set X, while the two outcomes are beyond the decision maker's expectation if he is currently at the node e_1. It is required that the behavioural rule at different nodes in the same information set should be identical. Thus, the decision problem of choosing the behavioural rule assigned to e_3 are different between being at node e_1 and being in the information set X. At node e_1, the optimal behavioural rule assigned to e_3 should be the one which maximises a weighted sum of the payoffs of outcome z_3 and z_4. Compared, the optimal behavioural rule at e_3 should be the one which maximises a weighted sum of the payoffs of outcome z_3, z_4, z_5 and z_6. Notice that the

[1] The subscript "IS" is the abbreviation of "Information Set". In our paper, a definition contains the abbreviation "IS" if it implies the viewpoint of a sophisticated decision maker who analyses decision problems with imperfect recall over information sets. A sophisticated decision maker realises he meets identical situation at any node in a non-singleton information set and would take the same action at each node in equilibrium. Thus, the sophisticated decision maker focuses on things happened at an information set rather than a particular node in a non-singleton information set.

behavioural rule determines the weights among payoffs of different outcomes. Thus, the optimal behavioural rule at e_3 is probably different between the occasion of being at node e_1 and that of being in the information set X. Thus, being at node e_1, it is different for the decision maker to think the behaviour D would lead himself to the node e_3 or the information set X. Then, it is questionable that what would be the decision maker's optimal strategy if he realises he is currently at node e_1 and he will not be able to distinguish between the two nodes e_3 and e_4 in his future arrival to the information set X.

We say the decision maker is **sophisticated** if he knows the structure of the decision problem, and when reconsidering at an information set, he is able to notice the fact that he definitely cannot distinguish among different nodes once he reaches a later information set with imperfect recall and does the reconsideration again at that information set. For example, there are two information sets X_1 and X_2 such that $X_1 \prec X_2$, the decision maker shows imperfect recall at information set X_2. In fact, under the partial order $\prec$, there is only one immediate predecessor information set for each information set. However, with the relation $\prec'$, there can be more than one immediate predecessor information set of each information set.

Definition 2.5.2 *Define $P'_{IS}(\cdot) : \mathscr{X} \cup Z \rightrightarrows \mathscr{X}$ the* ***immediate predecessor information set correspondence*** *over the relation $\prec'$. If $X_1 \in P_{IS}(X_2)$, then $\exists h_1 \in X_1, h_2 \in X_2$ such that $h_1 = P(h_2)$. If $P'_{IS}(\{z\}) = X$, then $\exists h \in X$ such that $P_{IS}(z) = h$.*

Notice that $P'_{IS}(z), z \in Z$ is actually a function instead of a correspondence. Assume that $X_1 \in P'_{IS}(X_2)$. It is reasonable that the decision maker takes his future indistinguishability at information set X_2 into consideration when he chooses an optimal strategy at information set X_1. In other words, the decision maker is called to be sophisticated if he interprets that his behaviour will lead him to an information set instead of a particular node in that information set. Then, being at X_1, the decision maker should consider not only the terminal nodes which can be reached by a completely mixed behavioural strategy starting from any node in X_1 but also the terminal nodes that can be reached by a completely mixed behavioural strategy starting from any node in X_2. There might be some nodes in X_2 which can never be reached by any behaviour from X_1. We still use the example of "forget previous acquired information" to explain the facts. If we assume $X_1 = \{e_1\}$ and $X_2 = X$, $e_4 \in X$ cannot be reached by any behaviour from $\{e_1\}$. z_5, z_6 can be reached by a completely mixed behavioural strategy starting from X but they cannot be reached by any strategy if the decision maker is currently at e_1. However, in our framework, the decision maker needs to consider terminal nodes z_5 and z_6 when he is at e_1. The reason behind it is that at $\{e_1\}$, the decision maker notices he might do the reconsideration again at X. Thus, when he reconsiders at $\{e_1\}$, as a sophisticated decision maker, he should be aware the potential occurrence of future reconsideration and evaluates the conditional expected payoffs of the terminal nodes z_5 and z_6 even the decision maker cannot reach them. It is different from the occasion of perfect recall, in which it does not matter whether we understand the decision maker's behaviour leads to a particular node or the information set containing that node.

We can also interpret the inconsistency of an optimal strategy before and during execution in the example of absentminded driver paradox (see Figure 3.1) is due to the different perspectives of a decision maker whether a behaviour leads the decision maker to a history or an information set. Alternatively, the inconsistency is due to whether the decision maker is sophisticated or not. The framework we propose in this paper is how a sophisticated decision maker chooses his optimal strategy during execution.

In the example of absentminded driver, the classic calculation rule implies the expected payoffs at information set I is the weighted sum of the expected payoffs at node e_1 and e_2. To the contrary, a sophisticated decision maker who is assumed to be currently at node e_1 and implement the behaviour C, should notice he would forget his past moves and be not able to confirm his current position once he arrives at node e_2. Similar to welfare symmetry proposed in Dimitri (1999, 2009), the decision maker's attitude to the conditional expected payoffs at any node from the same information set should be identical since he is indistinguishable with different nodes, and it should also be identical to his attitude to the expected payoffs of the whole information set. The idea of identical attitude should be incorporated into the calculation of expected payoffs at an information set. In this paper, we are going to calculate the conditionally expected payoffs in a recursive form while integrating with the idea of identical attitude. We call this method to calculate expected payoffs the sophisticated recursive calculation rule.

2.5.1 Two Examples

In this subsection, we calculate the conditional expected payoffs at node e_1 in the example of "forget previous acquired information" and the conditional expected payoffs at information set I in the example of absentminded driver paradox using the sophisticated recursive calculation rule. Then, we are going to discuss time consistency of an ex-ante optimal strategy in terms of the new calculation rule.

In the example of "forget previous acquired information", denote a behavioural strategy $\sigma = (\sigma_{e_0}, \sigma_{e_1}, \sigma_{e_2}, \sigma_X)$, in which σ_{e_0} is the probability assigned to the action G at node e_0, σ_{e_1} and σ_{e_2} are the probability assigned to the action D at node e_1 and e_2 respectively, and σ_X is the probability assigned to the action L at information set X. Then, the expected payoffs of a sophisticated decision maker at information set X is

$$\begin{aligned}
E^{IS}u(\sigma|X) &= \sigma_X(\mu(e_3|X,\sigma)u(z_3) + \mu(e_4|X,\sigma)u(z_4)) + (1-\sigma_X)(\mu(e_3|X,\sigma)u(z_5) + \mu(e_4|X,\sigma)u(z_6)) \\
&= \sigma_X(\frac{\sigma_{e_1}}{\sigma_{e_1}+\sigma_{e_2}} \cdot 5 + \frac{\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}} \cdot 0) + (1-\sigma_X)(\frac{\sigma_{e_1}}{\sigma_{e_1}+\sigma_{e_2}} \cdot 0 + \frac{\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}} \cdot 6) \\
&= \frac{5\sigma_{e_1} - 6\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}} \cdot \sigma_X + \frac{6\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}}.
\end{aligned}$$

Denote $\sigma^\star$ the optimal strategy of this example, $\sigma^\star = (1,0,1,0)$. At information set X, with the knowledge $(\sigma^\star_{e_1}, \sigma^\star_{e_2}) = (0,1)$, the expected payoffs is

$$E^{IS}u(\sigma_X|X,(\sigma^\star_{e_1}, \sigma^\star_{e_2})) = 6(1-\sigma_X).$$

Then, it is obvious that the optimal strategy which the sophisticated decision maker chooses at information set X is $(1,0,1,0)$, which is identical to the optimal strategy $\sigma^\star$.

Now, the decision maker at node e_1 reevaluates different strategies to choose an optimal one. In the classic model, the decision maker has an alternative better strategy $\sigma = (1,1,0,1)$ than the ex-ante optimal strategy $\sigma^\star$. If the decision maker calculates the expected payoffs at node e_1 in a sophisticated perspective, then it should be

$$\begin{aligned}
E^{IS}u(\sigma|\{e_1\}) &= (1-\sigma_{e_1})u(z_1) + \sigma_{e_1} \cdot \frac{p(X|\emptyset,\sigma)}{p_{\{e_1\}}(X|\emptyset,\sigma)} E^{IS}u(\sigma|X) \\
&= 2(1-\sigma_{e_1}) + 5\sigma_{e_1}\sigma_X + 6\sigma_{e_2}(1-\sigma_X)
\end{aligned}$$

At node e_1, the decision maker knows his ex-ante optimal strategy in which $\sigma^\star_{e_2} = 1$,then the expected payoffs can be written as

$$E^{IS}u((\sigma_{e_1}, \sigma_X)|\{e_1\}, \sigma^\star_{e_2}) = 2(1-\sigma_{e_1}) + 5\sigma_{e_1}\sigma_X + 6(1-\sigma_X)$$

Thus, the optimal strategy conditional at the node e_1 is still $(1,0,1,0)$ which is identical to the ex-ante optimal strategy. It is different from the conclusion that if we calculate the conditional expected payoffs in the classic method, in which the decision maker has an alternative better strategy at node e_1. It shows the time consistency of the ex-ante optimal strategy in a sophisticated perspective. Denote $p_{\{e_1\}}(X|\emptyset, \sigma)$ the probability that the decision maker reaches the information set X via one move from the node e_1, i.e., $p_{\{e_1\}}(X|\emptyset, \sigma) = p(e_3|\emptyset, \sigma)$. The portion $p(X|\emptyset, \sigma)/p_{\{e_1\}}(X|\emptyset, \sigma)$ in the expected payoff function is a parameter which is relevant to the decision maker's mental evaluation of the expected payoff of information set X conditional being at the node e_1.

Another example is the absentminded driver paradox. Notice that the probability of choosing C at an intersection is denoted by p When the driver goes onto the highway, i.e. the decision maker is at the information set I, the expected payoffs should be

$$\begin{aligned} E^{IS}u(p|I) &= \mu(e_1|I,p)[(1-p)u(z_1) + pE^{IS}u(p|I)] + \mu(e_2|I,p)[pu(z_3) + (1-p)u(z_2)] \\ &= \frac{p}{p+1}E^{IS}u(p|I) + \frac{p}{p+1}(p + 4(1-p)) \end{aligned}$$

Then, the expected payoffs are

$$E^{IS}u(p|I) = p^2 + 4p(1-p).$$

It is obvious that $E^{IS}u(p|I) = Eu(p|\emptyset)$, in which $Eu(p|\emptyset)$ is the expected payoffs calculated before the decision problem starts. Thus, the ex-ante optimal strategy $p^\star = 2/3$ is also the optimal strategy conditional on the information set I in a sophisticated perspective.

We can see from the calculation results of these two examples that the ex-ante optimal strategy is time consistent under the construction of the new calculation rule on expected payoffs at an information set. In the next subsection, we will construct the calculation rule of the expected payoffs at an information set in a sophisticated perspective.

2.5.2 Calculation Rules Applied to Decision Problems with Perfect Recall

Firstly, we would like to rewrite the classic expected payoff function in a recursive form. The expected payoff function at information set X is

$$Eu(\sigma|X) = \sum_{h\in X} \mu(h|X,\sigma) \sum_{z\in Z(h)} p(z|h,\sigma)u(z). \tag{2.1}$$

In decision problems with perfect recall, $\mu(h|X,\sigma)$ does not dependent on σ. We can substitute $\rho(h|X)$ for $\mu(h|X,\sigma)$ with the presence of perfect recall. In other words, the non-singleton information set in decision problems with perfect recall is caused by unknown moves by the chance player. Thus, the expected payoff function is

$$\begin{aligned}Eu(\sigma|X) &= \sum_{h\in X} \rho(h|X) \sum_{z\in Z(h)} p(z|h,\sigma)u(z)\\ &= \sum_{h\in X} \rho(h|X) \sum_{a\in A(X)} p((h,a)|h,\sigma) \sum_{z\in Z((h,a))} p(z|(h,a),\sigma)u(z).\end{aligned}$$

Assume that the decision maker moves at (h,a), whenever $h \in X$ and $a \in A(X)$. Then, for $h_1,...,h_k \in X$ and $a_1,...,a_k \in A(X)$, assume that the nodes (h_1,a_1), (h_2,a_2),..., (h_k,a_k) belong to an information set and for simplicity, the rest (h,a) which is different from $(h_t,a_t), t = \{1,...,k\}$ respectively belongs to a different singleton information set $\{(h,a)\}$ or $X_{(h,a)}$. Denote $X_{(h,a)}$ the information set containing the node (h,a). Here, we use the history of a node to represent the node. In other words, for any two nodes (h,a) and (h',a') which are different from $(h_t,a_t), t = \{1,...,k\}$, we have $X_{(h,a)} \neq X_{(h',a')} \neq X_{(h_1,a_1)}$. To satisfy the requirement of perfect recall, $a_1 = a_2 = ... = a_k$. For those $(h,a) \notin X_{(h_1,a_1)}$, $Z(X_{(h,a)}) = Z((h,a))$. The expected payoff function at node (h,a) is

$$\begin{aligned}Eu(\sigma|X_{(h,a)}) &= Eu(\sigma|(h,a))\\ &= \sum_{z\in Z((h,a))} p(z|(h,a),\sigma)u(z) \quad = \sum_{z\in Z\left(X_{(h,a)}\right)} p(z|(h,a),\sigma)u(z)\end{aligned}$$

For nodes (h_1,a_1), (h_2,a_2),...,(h_k,a_k) which are in the same information set. The expected payoff function at information set $X_{(h_1,a_1)}$ is

$$\begin{aligned}Eu(\sigma|X_{(h_1,a_1)}) &= \sum_{i=1}^{k} \mu((h_i,a_i)|X_{(h_1,a_1)}) \sum_{z\in Z((h_i,a_i))} p(z|(h_i,a_i),\sigma)u(z)\\ &= \sum_{i=1}^{k} \rho((h_i,a_i)|X_{(h_1,a_1)}) \sum_{z\in Z((h_i,a_i))} p(z|(h_i,a_i),\sigma)u(z).\end{aligned}$$

Then, the expected payoff function at information set X is

$$
\begin{aligned}
Eu(\sigma|X) &= \sum_{h\in X}\rho(h|X)\sum_{a\in A(X)}p((h,a)|h,\sigma)\sum_{z\in Z((h,a))}p(z|(h,a),\sigma)u(z) \\
&= \sum_{\substack{h\in X, a\in A(X)\\(h,a)\notin X_{(h_1,a_1)}}}\rho(h|X)p((h,a)|h,\sigma)\sum_{z\in Z((h,a))}p(z|(h,a),\sigma)u(z) \\
&+ \sum_{\substack{h\in X, a\in A(X)\\(h,a)\in X_{(h_1,a_1)}}}\rho(h|X)p((h,a)|h,\sigma)\sum_{z\in Z((h,a))}p(z|(h,a),\sigma)u(z) \\
&= \sum_{\substack{h\in X, a\in A(X)\\(h,a)\notin X_{(h_1,a_1)}}}\rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)}) \\
&+ \sum_{i=1}^{k}\rho(h_i|X)p((h_i,a_i)|h_i,\sigma)\sum_{z\in Z((h_i,a_i))}p(z|(h_i,a_i),\sigma)u(z)
\end{aligned}
$$

We have analysed that $a_1 = a_2 = ... = a_k$, then $p((h_1,a_1)|h_1,\sigma) = p((h_2,a_2)|h_2,\sigma) = ... = p((h_k,a_k)|h_k,\sigma)$. Thus,

$$\rho((h_i,a_i)|X_{(h_1,a_1)}) = \frac{p((h_i,a_i)|\sigma)}{\sum_{j=1}^{k}p((h_j,a_j)|\sigma)} = \frac{p(h_i|\sigma)}{\sum_{j=1}^{k}p(h_j|\sigma)},$$

We transform the equation and then get,

$$\rho(h_i|X) = \sum_{j=1}^{k}\rho(h_j|X)\rho((h_i,a_i)|X_{(h_1,a_1)}).$$

Then, substitute the equation into $E(\sigma|X)$,

$$
\begin{aligned}
Eu(\sigma|X) = &\sum_{\substack{h\in X, a\in A(X)\\ (h,a)\notin X_{(h_1,a_1)}}} \rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)})\\
&+\sum_{i=1}^{k}\sum_{j=1}^{k}\rho(h_j|X)\rho((h_i,a_i)|X_{(h_1,a_1)})p((h_i,a_i)|h_i,\sigma)\sum_{z\in Z((h_i,a_i))}p(z|(h_i,a_i),\sigma)u(z)\\
=&\sum_{\substack{h\in X, a\in A(X)\\ (h,a)\notin X_{(h_1,a_1)}}} \rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)})\\
&+\sum_{j=1}^{k}\rho(h_j|X)p((h_j,a_j)|h_j,\sigma)\sum_{i=1}^{k}\rho((h_i,a_i)|X_{(h_1,a_1)})\sum_{z\in Z((h_i,a_i))}p(z|(h_i,a_i),\sigma)u(z)\\
=&\sum_{\substack{h\in X, a\in A(X)\\ (h,a)\notin X_{(h_1,a_1)}}} \rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)})\\
&+\sum_{j=1}^{k}\rho(h_j|X)p((h_j,a_j)|h_j,\sigma)Eu(\sigma|X_{(h_1,a_1)})\\
=&\sum_{h\in X, a\in A(X)} \rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)})
\end{aligned}
$$

Thus, we can rewrite the conditional expected payoffs at an information set in recursive form, which is

$$
Eu(\sigma|X) = \sum_{h\in X, a\in A(X)} \rho(h|X)p((h,a)|h,\sigma)Eu(\sigma|X_{(h,a)}) \tag{2.2}
$$

Now assume that the chance player moves at some (h,a), $h\in X$ and $a\in A(X)$. For simplicity, 1) chance player moves at most one step at any (h,a); 2) all the information sets containing (h,a) or (h,a,c) (if chance player moves at (h,a)) are singletons except for one information set. Denote that non-singleton information set X', $X' = \{(h_1,a_1),...,(h_l,a_l),(h_{l+1},a_{l+1},c_{l+1}),...,(h_k,a_k,c_k)\}$, i.e., the decision maker moves at l nodes in X' and chance player moves at the other $k-l$ nodes. There might exist (h_i,a_i,c_i) and (h_j,a_j,c_j) such that $h_i = h_j$, $a_i = a_j$, but $c_i \neq c_j$. Alternatively, we could rewrite every node in the form of (h,a,c). For those (h,a), at which the decision maker moves, c is assumed to be the deterministic action of chance player. Then, the expected payoff function is

$$
\begin{aligned}
Eu(\sigma|X) &= \sum_{h\in X} \rho(h|X) \sum_{z\in Z(h)} p(z|h,\sigma)u(z) \\
&= \sum_{h\in X} \rho(h|X) \sum_{\substack{a\in A(X) \\ c\in A((h,a))}} p((h,a,c)|h,\sigma) \sum_{z\in Z((h,a,c))} p(z|(h,a),\sigma)u(z) \\
&= \sum_{\substack{h\in X, a\in A(X), \\ c\in A((h,a)), (h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma) \sum_{z\in Z((h,a,c))} p(z|(h,a,c),\sigma)u(z) \\
&+ \sum_{\substack{h\in X, a\in A(X), \\ c\in A((h,a)), (h,a,c)\in X'}} \rho(h|X)p((h,a,c)|h,\sigma) \sum_{z\in Z((h,a))} p(z|(h,a,c),\sigma)u(z) \\
&= \sum_{\substack{h\in X, a\in A(X), \\ c\in A((h,a)), (h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
&+ \sum_{i=1}^{k} \rho(h_i|X)p((h_i,a_i,c_i)|h_i,\sigma) \sum_{z\in Z((h_i,a_i,c_i))} p(z|(h_i,a_i,c_i),\sigma)u(z)
\end{aligned}
$$

The expected payoff function at the information set X' is

$$
Eu(\sigma|X') = \sum_{i=1}^{k} \rho((h_i,a_i,c_i)|X') \sum_{z\in Z((h_i,a_i,c_i))} p(z|(h_i,a_i,c_i),\sigma)u(z).
$$

It is easy to know that,

$$
\begin{aligned}
\rho(h_i|X) &= \frac{p(h_i|\sigma)}{\sum_{h\in X} p(h|\sigma)} \\
&= \frac{p(h_i|\sigma)\rho(c_i)}{\sum_{j=1}^{k} p(h_j|\sigma)\rho(c_j)} \cdot \frac{\sum_{j=1}^{k} p(h^{k}_{j=1} p(h_j|\sigma)\rho(c_j)}{\rho(c_i)\sum_{h\in X} p(h|\sigma)} \\
&= \frac{\sum_{j=1}^{k} p(h_j|\sigma)\rho(c_j)}{\rho(c_i)\sum_{h\in X} p(h|\sigma)} \cdot \frac{p((h_i,a_i,c_i)|\sigma)}{\sum_{j=1}^{k} p((h_j,a_j,c_j)|\sigma)} \\
&= \frac{\sum_{j=1}^{k} p(h_j|\sigma)\rho(c_j)}{\rho(c_i)\sum_{h\in X} p(h|\sigma)} \cdot \rho((h_i,a_i,c_i)|X') \\
&= \sum_{j=1}^{k} \frac{\rho(c_j)}{\rho(c_i)} \cdot \rho(h_j|X) \cdot \rho((h_i,a_i,c_i)|X')
\end{aligned}
$$

Then, substitute the equation into the expected payoff function $Eu(\sigma|X)$,

$$
\begin{aligned}
Eu(\sigma|X) = & \sum_{\substack{h\in X, a\in A(X),\\ c\in A((h,a)),(h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
& + \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{\rho(c_j)}{\rho(c_i)}\cdot\rho(h_j|X)\cdot\rho((h_i,a_i,c_i)|X')p((h_i,a_i,c_i)|h_i,\sigma)\sum_{z\in Z((h_i,a_i,c_i))}p(z|(h_i,a_i,c_i),\sigma)u(z) \\
= & \sum_{\substack{h\in X, a\in A(X)\\ c\in A((h,a)),(h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
& + \sum_{j=1}^{k}\rho(c_j)\rho(h_j|X)\sum_{i=1}^{k}\rho((h_i,a_i,c_i)|X')\cdot\frac{p((h_i,a_i,c_i)|h_i,\sigma)}{\rho(c_i)}\sum_{z\in Z((h_i,a_i,c_i))}p(z|(h_i,a_i,c_i),\sigma)u(z) \\
= & \sum_{\substack{h\in X, a\in A(X),\\ c\in A((h,a)),(h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
& + \sum_{j=1}^{k}\rho(c_j)\rho(h_j|X)p((h_j,a_j)|h_j,\sigma)\sum_{i=1}^{k}\rho((h_i,a_i,c_i)|X')\sum_{z\in Z((h_i,a_i,c_i))}p(z|(h_i,a_i,c_i),\sigma)u(z) \\
= & \sum_{\substack{h\in X, a\in A(X),\\ c\in A((h,a)),(h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
& + \sum_{j=1}^{k}\rho(h_j|X)p((h_j,a_j,c_j)|h_j,\sigma)\sum_{i=1}^{k}\rho((h_i,a_i,c_i)|X')\sum_{z\in Z((h_i,a_i,c_i))}p(z|(h_i,a_i,c_i),\sigma)u(z) \\
= & \sum_{\substack{h\in X, a\in A(X),\\ c\in A((h,a)),(h,a,c)\notin X'}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)}) \\
& + \sum_{j=1}^{k}\rho(h_j|X)p((h_j,a_j,c_j)|h_j,\sigma)Eu(\sigma|X') \\
= & \sum_{\substack{h\in X, a\in A(X),\\ c\in A((h,a))}} \rho(h|X)p((h,a,c)|h,\sigma)Eu(\sigma|X_{(h,a,c)})
\end{aligned}
$$

Therefore, we can see from the calculation that, the expected payoff function can be rewritten in a recursive form even chance player participates.

It has been displayed that, equation 2.1 is equivalent to equation 2.2 with the assumption of perfect recall. We call the calculation rule in terms of terminal nodes, i.e. equation 2.1, the traditional calculation rule and the calculation rule relates to the expected payoffs at later information sets, i.e. equation 2.2, the recursive calculation rule.

2.5.3 Sophisticated Recursive Calculation Rule

The equivalence between the traditional calculation rule and the recursive calculation rule of the expected payoffs is not valid anymore with the presence of imperfect recall. Take “forget previous acquired information" as an example (see Figure 3.5), the expected payoff function at information set is

$$\begin{aligned} Eu(\sigma|X) &= \sigma_X(\mu(e_3|X,\sigma)u(z_3)+\mu(e_4|X,\sigma)u(z_4))+(1-\sigma_X)(\mu(e_3|X,\sigma)u(z_5)+\mu(e_4|X,\sigma)u(z_6)) \\ &= \frac{5\sigma_{e_1}-6\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}}\cdot\sigma_X+\frac{6\sigma_{e_2}}{\sigma_{e_1}+\sigma_{e_2}}. \end{aligned}$$

In the meanwhile, the expected payoff function at node e_1 is

$$Eu(\sigma|\{e_1\}) = 2(1-\sigma_{e_1})+5\sigma_{e_1}\sigma_X.$$

It is obvious that

$$Eu(\sigma|\{e_1\}) \neq 2(1-\sigma_{e_1})+\sigma_{e_1}E(\sigma|X).$$

In this case, we could officially modify the recursive calculation rule and officially develop the sophisticated recursive calculation rule to analyse decision problems with imperfect recall.

By recursion, at information set X, the decision maker needs to consider all the terminal nodes that he considers at an immediate successor information set X' of information set X. It is possible that X and X' have a common immediate successor information set. We use an abstract decision problem to explain it. As shown in Figure 3.8, $X_1 \in P'_{IS}(X_2)$ and $X_1, X_2 \in P'_{IS}(X_3)$, thus X_3 is the common immediate successor information set of X_1, X_2. The decision maker might consider some terminal nodes in more than one IS-path, i.e., evaluate those terminal nodes more than once.

Definition 2.5.3 *We say an* ***IS-path*** *from information set X to X' (or $\{z\}$) exists if there are several information sets $X'_1,...,X'_k$ such that $X \in P'_{IS}(X'_1), X'_1 \in P'_{IS}(X'_2)$,..., $X'_k \in P'_{IS}(X')$ (or $\exists h \in X'_k$ such that $(h,a)=z$, $a \in X_k$ if there is an IS-path from X to X' (or $\{z\}$).*

If the decision maker is at X_3, he should considers the terminal nodes z_i, $i=\{6,7,8,9\}$. If the decision maker is at X_2, he should considers the terminal nodes z_3, z_4, z_5 and all the nodes considered at X_3 since $X_2 \in P'_{IS}(X_3)$.

The decision maker considers z_i, $i = \{3,4,5,6,7,8,9\}$ at X_2. Then, when the decision maker is at information set X_1, he should evaluate the terminal nodes z_i, $i = \{6,7,8,9\}$ through two different IS-paths, one passes information set X_2 and one does not. However, we know that there is only one path from the start of the decision problem to any terminal node. Thus, based on the fact of being at information set X_1, there is at most one IS-path which can be realised if there are more than one IS-path to a terminal node. For any terminal node such that there are more than one IS-path from the information set where the decision maker is at to the terminal node, a decision maker either cannot reach that terminal node, or reaches that terminal node through the information sets belonging to one IS-path.

Now, we discuss back to the inspiration behind the calculation of conditional expected payoffs. A decision maker maximises a weighted sum of the payoffs assigned at different terminal nodes, in which the weight is the conditional probability from his current node to each terminal node. However, in the presence of imperfect recall, a decision maker does not know which node he is currently located in the decision problem tree. Or the decision maker knows he will be ignorant about the node he is located at that time as the decision problem executes. In conventional decision theory, the decision maker evaluates the decision problem in the same way no matter whether he knows the fact of future imperfect recall or not except for the restriction of available strategies. We would like to develop a new way of evaluation which makes it different between the situation that the decision maker knows he will present imperfect recall and the situation that he does not know.

We propose that a sophisticated decision maker should take the fact of future imperfect recall into consideration when he chooses a strategy. Some of the outcomes that will never occur conditional at the current information set should also be considered currently. It is because the decision maker knows the structure of the decision problem and the fact he will lose the information of the current history he is located if the information set reached by his one-shot move presenting imperfect recall. The outcomes considered at the information sets he thinks he might reach in the future due to imperfect recall should also be considered even though the current decision maker knows they will never be reached by any completely mixed behavioural strategy. Therefore, in our approach, a sophisticated decision maker should consider the choice of a conditionally optimal strategy at an information set X with the viewpoint that each move leads him to an information set instead of a particular node. When estimating the conditional probability of reaching a terminal node, we actually calculate the probability of reaching a terminal node from the current information set through IS-paths. By recursion, the probability of an IS-path is the multiplication of the modified probability of reaching an information set X_2 from its immediate processor information set X_1 by one step move, where X_1, X_2 belong to the IS-path involved. The modification represents the decision maker's attitude to the expected payoffs at information set X_2 when he is at X_1.

If there is more than one IS-path to a terminal node, there must be at least one information set in each path such that the decision maker does not know whether he can reach that terminal node by a completely mixed behavioural strategy. It is because he loses the information about the node he is currently located caused by his forgetfulness. The decision maker takes his future forgetfulness into consideration. He chooses a strategy as if he cannot distinguish which IS-path contains the real path. It is possible that the decision maker cannot reach a terminal node by any completely mixed behavioural strategy when there exists an IS-path from the decision maker's current information set to that terminal node. For example, as shown in Figure 3.8, the decision maker at node d_5 knows he cannot reach the terminal node z_8 by any completely behavioural strategy despite there is an IS-path from information set X_2 to the set containing the terminal node z_8, $\{z_8\}$. The decision maker does not know whether he is at node d_5 or d_6 currently. He does not know whether he will move to the information set X_3 although he knows he will be at node d_7 if he moves to information set X_3. However, once he has moved to information set X_3, he would not know whether he is at node d_7 or d_8, terminal nodes z_8 and z_9 are also potential outcomes of the decision maker at information set X_3. Then, as a sophisticated decision maker at information set X_2, he should be able to imagine the decision problem he would face once he moves to information set X_3. He should also know the fact that his behaviour at the current information set X_2 might influence his evaluation of the decision problem he faces at information set X_3 in the future through the belief system. Thus, he should also take the terminal nodes z_8 and z_9 into consideration although he clearly knows he will never reach those two nodes. Since the decision maker acts as if he cannot figure out which IS-path contains the real path to a terminal node when there is more than one IS-paths, the decision maker is assumed to assign equal probability to each one. Thus, when calculating the conditional expected payoff function of a sophisticated decision maker, the weight of payoffs at each terminal node is the average of attitude modified probabilities of IS-paths from the current information set to each terminal node.

We would like to calculate the expected payoff function conditional of a sophisticated decision maker at an information set by a recursive calculation rule. We call it **sophisticated recursive calculation rule**. It contains two steps. The first step is to calculate the average of attitude modified probabilities of IS-paths from the current information set where the decision maker is to each terminal node. In the second step, we substitute the payoffs assigned at a terminal node for the terminal node. We give several notations before the description of sophisticated recursive calculation rule.

Notation 2.5.2 *Denote* $\mathscr{Y} := \mathscr{X} \cup (\cup_{z \in Z}\{z\})$ *and* $Y \in \mathscr{Y}$ *can be an information set* X *or a set containing a terminal node* z*. Denote* Y_h *the information set containing* h *or the set containing a terminal node* z *if* $h = z$*.*

Additionally, we extend the relation $\prec'$ to $\mathscr{Y}$, we say $X \prec' \{z\}$ if there is a history $h \in X$ such that h is a subhistory of z.

Notation 2.5.3 *Denote* $z \in Z_{IS}(X)$ *if there is a IS-path from the information set* X *to the terminal node* z*.*

Notation 2.5.4 *Denote* $p_X(Y|\emptyset, \sigma)$ *the sum of probabilities of reaching those nodes* h *satisfying:* $h \in \hat{Y}$*, and* $\exists x \in X$ *such that* $P(h) = x$*, given the strategy* σ*. Then,*

$$p_X(Y|\emptyset,\sigma) := \sum_{\substack{h' \in X, a' \in A(X) \\ (h',a') \in Y \neq X}} p((h',a')|\emptyset,\sigma),$$

We say an IS-path from X to the set of terminal node Z, $\{z\}$, is **feasible** with a behavioural strategy σ, if for any $X' \in \mathscr{X}, Y' \in \mathscr{Y}$ on the IS-path such that $X' \in P_{IS}(Y')$, we have $p_{X'}(Y'|\emptyset,\sigma) > 0$.

Notation 2.5.5 *Denote* $z \in Z_{IS}(X, \sigma)$*, if there exists a feasible IS-path from* X *to* $\{z\}$ *with the behavioural strategy* σ*.*

Then, we would like to define a function which specifies the average modified probability of reaching every terminal node through IS-paths to that terminal node. The modification illustrates how the decision maker consider about his one-shot move to another information set (the one-shot move could lead to the same information set if the decision maker is at an information set with absentmindedness).

Definition 2.5.4 *Define* $\Phi(\cdot|X):\Sigma\to\times_{|Z|}(\Re\times Z)$ *the* ***attitude modified IS-path probability function*** *conditional on being at information set X, where* $|Z|$ *is the number of terminal nodes. To simplify the function form, a terminal node is omitted in the function if there is no feasible IS-path from the information set X to that terminal node. If there is no chance player in the decision problem, for a behavioural strategy* σ*,*

$$\Phi(\sigma|X)=f_{X,\sigma}\left(\sum_{a\in A(X)}\sigma_X(a)\sum_{h\in X}\mu(h|X,\sigma)\cdot\lambda(h,a|\sigma)\cdot\Phi(\sigma|Y_{(h,a)})\right),$$

where

$$\lambda(h,a|\sigma):=\begin{cases}1 & \text{if } h\in Y_{(h,a)} \text{ or } (h,a)\in Z\\ 0 & \text{if } h\notin Y_{(h,a)}, p_{X_h}(Y_{(h,a)}|\emptyset,\sigma)=0\\ p(Y_{(h,a)}|\emptyset,\sigma)/p_{X_h}(Y_{(h,a)}|\emptyset,\sigma) & \text{if } h\notin Y_{(h,a)}, p_{X_h}(Y_{(h,a)}|\emptyset,\sigma)\neq 0\end{cases}$$

and

$$f_{X,\sigma}(\cdot)=\left(1-\prod_{z'\in Z_{IS}(X,\sigma)}1_{\{z'\}}(\cdot)\right)\sum_{z\in Z_{IS}(X,\sigma)}\frac{1_{\{z\}}(\cdot)\cdot z}{\sum_{\substack{X\in P'_{IS}(Y),Y\in\mathscr{Y},\\ p_X(Y|\emptyset,\sigma)>0}}1_{Z_{IS}(Y,\sigma)}(z)}$$
$$+\left(\prod_{z'\in Z_{IS}(X,\sigma)}1_{\{z'\}}(\cdot)\right)\Phi(\sigma|X)$$

Notice that $\Phi(\sigma|Y_h)=z$ and $Z_{IS}(Y_h)=\{z\}$ if $h=z$, $\forall\sigma\in\Sigma$. It is explained as that the probability for a decision maker to reach the terminal node z through IS-paths is 1 if he is at the terminal node z. The set of terminal nodes which can be reached by an IS-path from the terminal node z is $\{z\}$.

Definition 2.5.5 *Define* $f_{X,\sigma}(\cdot):\times_{|Z|}(\Re\times Z)\to\times_{|Z|}(\Re\times Z)$ *the* ***average function*** *which assigns equal probability to each feasible IS-path from X to the set containing the same terminal node with the behavioural strategy* σ*.*

Notice that the average function separates the situation that his immediate move lead him to the same information set from the situation that he moves to a different Y. If a behavioural strategy is completely mixed, $Z_{IS}(X) = Z_{IS}(X,\sigma)$ for any information set X.

The λ in the equation shows the decision maker's attitude of expected payoffs of information set $X_{(h,a)}$ or the payoffs at the terminal node $z = (h,a)$ conditional being at information set X. Being at information set X_h, we could see the fraction $1/\lambda$ of $Y_{(h,a)}$. As we have stated, at any node in an information set, the decision maker's current evaluation of the decision problem is identical to the expected payoffs of that information set. When the decision maker evaluates the expected payoffs at information set X_h, he would consider the expected payoffs at $Y_{(h,a)}$ as $\Phi(\sigma|Y_{(h,a)})/(1/\lambda) = \lambda \cdot \Phi(\sigma|Y_{(h,a)})$. Thus, the attitude modified IS-path probability function indicates that, conditional being at information set X_h, the decision maker has probability $\mu(h|X,\sigma)$ of being at history h. Based on being at history h, his strategy assigns probability $\sigma_X(a)$ to reach the node (h,a), and reaches $Y_{(h,a)}$. His evaluation of the expected payoffs at $Y_{(h,a)}$ is $\lambda \cdot \Phi(\sigma|Y_{(h,a)})$. Therefore, the sum over all the combinations (h',a') such that $h' \in X_h, a' \in A(X)$ makes up the expected payoff function at information set X.

The function $1_K(\cdot)$ where $K \in \mathscr{X} \cup Z$ is a characteristic function. The denominator

$$\sum_{\substack{X \in P'_{IS}(Y), Y \in \mathscr{Y} \\ p_X(Y|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y,\sigma)}(z)$$

is the number of IS-paths from information set X to the terminal node z. Thus, the function assigns equal probability to each feasible IS-path to the same information set.

Now the chance player works in the decision problem. For simplicity, assume that the chance player moves at (h_1,a_1) where $h_1 \in X, a_1 \in A(X)$ and $c^1 \in A((h_1,a_1))$, for other $(h,a) \neq (h_1,a_1)$ such that $h \in X, a \in A(X)$, the decision maker works. the calculation rule becomes

$$\Phi(\sigma|X,c) = f_{X,\sigma}[\sum_{\substack{a \in A(X), h \in X \\ (h,a) \neq (h_1,a_1)}} \sigma_X(a)\mu(h|X,\sigma,c)\lambda(h,a|\sigma,c)\Phi(\sigma|Y_{(h,a)},c)$$
$$+ \sigma_X(a_1)\mu(h_1|X,\sigma,c) \sum_{c^1 \in A((h_1,a_1))} \rho(c^1)\lambda(h_1,a_1,c^1|\sigma,c)\Phi(\sigma|Y_{(h_1,a_1,c^1)},c)]$$

where

$$\lambda(h_1,a_1,c^1|\sigma,c) := \begin{cases} 1 & \text{if } h_1 \in Y_{(h_1,a_1,c^1)} \text{ or } (h_1,a_1,c^1) \in Z \\ 0 & \text{if } h_1 \notin Y_{(h_1,a_1,c^1)}, p_{Y_{h_1}}(Y_{(h_1,a_1,c^1)}|\emptyset,\sigma,c) = 0 \\ \frac{p(Y_{(h_1,a_1,c^1)}|\emptyset,\sigma,c)}{p_{Y_{h_1}}(Y_{(h_1,a_1,c^1)}|\emptyset,\sigma,c)} & \text{if } h_1 \notin Y_{(h_1,a_1,c^1)}, p_{X_{h_1}}(Y_{(h_1,a_1,c^1)}|\emptyset,\sigma,c) \neq 0 \end{cases}$$

and

$$p_{X_{h_1}}(Y_{(h_1,a_1,c^1)}|\emptyset,\sigma,c) = \sum_{\substack{h' \in X_{h_1}, a' \in A(X) \\ (h',a') \in Y_{(h_1,a_1,c^1)}}} p((h',a')|\emptyset,\sigma,c) + \sum_{c' \in A((h_1,a_1))} p((h_1,a_1,c')|\emptyset,\sigma,c).$$

In fact, we can see that there is no much difference even we take account of the chance player.

In the second step, we substitute the payoff of a terminal node for that node.

Definition 2.5.6 *Define $g(\cdot) : \times_{|Z|}(\Re \times Z) \to \Re$ the **substitution function**.*

$$g(\cdot) = \sum_{z \in Z} u(z) 1_{\{z\}}(\cdot).$$

Then apply the substitution function g to the attitude modified IS-path probability function Φ to calculate decision maker's conditional expected payoffs.

Definition 2.5.7 *Define the expected payoff function resulted from the sophisticated recursive calculation rule **IS expected payoff function**, denoted by $E^{IS}u$, then*

$$E^{IS}u(\sigma|X) = g\left(\Phi(\sigma|X)\right).$$

Result 2.5.1 *Assume that the decision maker implements the behavioural strategy σ and reconsiders at information set X with the belief system μ. Then, the attitude modifies IS-path probability function is*

$$\Phi(\sigma|X) = \frac{1}{P(X|\emptyset,\sigma)} \sum_{z \in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma) z,$$

and its corresponding IS expected payoff function is

$$E^{IS}u(\sigma|X) = \frac{1}{P(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)u(z).$$

Proof of Result 2.5.1. See the Appendix.

Result 2.5.2 *Assume that the decision maker implements a completely mixed behavioural strategy σ and reconsiders at information set X with the belief system μ. Then, the attitude modifies IS-path probability function is*

$$\Phi(\sigma|X) = \frac{1}{P(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X)} p(z|\emptyset,\sigma)z,$$

and the expected payoffs in terms of the sophisticated recursive calculation rule is

$$E^{IS}u(\sigma|X) = \frac{1}{P(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X)} p(z|\emptyset,\sigma)u(z).$$

Now we define the multi-agent form of a decision problem, in which a different agent is assigned to a set of information sets. The information sets in the same set are connected by at least an IS-path. A agent can control the behaviours at the information set included in the set of information sets which he is assigned to. If the decision maker plays optimally at any information set according to IS expected payoff function, the strategy that the decision maker plays is a common equilibrium of different multi-agent formations of the decision problem. As shown in Figure 3.9, each black point represents an information set, a circle including black points represents an agent. The black line between two information sets represents the decision maker could reach an information set by one-step move from the other information set. Assume that, information sets T,G,M,W are all the immediate successor information sets of the information set X. The information sets which are in the same blue circle as the information set T (G,M,W) are those to which there is at least an IS-path from T (G,M,W). Then, when the decision maker evaluates at X, the expected payoff assigned at T (G,M,W) is the one when the agent including T (G,M,W) plays one of his best response given any strategy at other information sets except those included in the set containing T (G,M,W). The decision problems we discuss in this paper are finite. We could calculate the expected payoffs at each end of the decision problem and then take them for granted when calculating the expected payoffs at earlier information sets. It is consistent with the inspiration of backward induction. A conditionally optimal strategy chosen at each information set at least satisfies one-shot deviation principle, and extra requirements of optimality. We could regard a strategy which is conditionally optimal at any information set according to the calculation rule satisfies a strong version of one-shot deviation principle.

2.5.4 Alternative Formation of the Sophisticated Recursive Calculation Rule

In the previous subsection, we use characteristic function to do the step of substitution. Alternatively, we can also explain the sophisticated recursive calculation rule in terms of matrices.

Since the decision problems we discuss are finite, we can use different $n \times 1$ matrices to express different terminal nodes separately if there are n terminal nodes in the decision problem we study. For example, if the terminal nodes are $z_k, k = \{1, ..., n\}$, we can define $z_k = (0, .., 0, 1, 0, ..., 0)^T$, $k = \{1, ..., n\}$. Every element is 0 in the $n \times 1$ matrix except the kth element which is 1. We substitute the matrix for the corresponding terminal node in the attitude modified IS-path probability function. Denote the new probability function $\Phi'(\cdot|X) : \Sigma \to \Re^{n\times 1}$. Furthermore, we can use a $n \times n$ matrix to express the equal probability when there is more than one IS-path from X to a terminal node and the substitution of payoffs. Then, the part which indicates the IS-path number in function $f'_X(\cdot) : \Re^{n\times n} \to \Re^{n\times 1}$ is

$$\sum_{z\in Z_{IS}(X,\sigma)} \frac{1_{\{z\}}(\cdot)\cdot z}{\sum_{\substack{X\in P'_{IS}(Y), Y\in\mathscr{Y},\\ p_X(Y|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y,\sigma)}(z)}.$$

It is written as

$$\begin{bmatrix} \frac{1}{\sum_{\substack{X\in P'_{IS}(Y), Y\in\mathscr{Y}\\ p_X(Y|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y)}(z_1)} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sum_{\substack{X\in P'_{IS}(Y), Y\in\mathscr{Y}\\ p_X(Y|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y)}(z_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{1}{\sum_{\substack{X\in P'_{IS}(Y), Y\in\mathscr{Y}\\ p_X(Y|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y)}(z_n)} \end{bmatrix}.$$

The new substitution function $g' : \Re^{n\times 1} \to \Re$ is

$$g'(\Phi) = (u(z_1), u(z_2), \cdots, u(z_n))\Phi.$$

It is easily to know that $E^{IS}u(\sigma|X) = g'(\Phi'(\sigma|X))$.

2.5.5 Cross-branch Information Sets

In the relation $\prec'$, there might exist two different information sets X and X' in a finite decision problem with imperfect recall, where $X \prec' X'$ and $X' \prec' X$, such as Gilboa formation of absentminded driver paradox (see Figure 3.7). We call such two information sets are cross-branch information sets. Gilboa's formation separates the decision maker into two selves, thus, transforms the absentminded driver paradox into a multiself game form. However, in our discussion about the cross-branch information sets, we only adopt the structure of Gilboa's formation but still model it in a one-self approach.

Definition 2.5.8 *If there are more than two information sets $X_1, ..., X_k$ such that $\forall i, j \in \{1, ..., k\}$, $X_i \prec' X_j$ and $X_j \prec' X_i$, then we call $X_1, ..., X_k$ are* ***cross-branch information sets****.*

Assume that the decision maker is at information set X_1, he does not know whether he has passed information set X_2 which indicates he is at node e_4 or he may reach information set X_2 after his immediate action which indicates he is at node e_1. We have two choices to deal with similar issues. The first one is the same as multiself approach. Although $X_1 \prec' X_2$, the decision maker cannot change to any other actions at information set X_2 but following the strategy chosen before arriving X_1 and X_2 since he potentially has passed it. Thus, the decision maker chooses a conditional optimal strategy based on given behavioural rule at X_2. Denote p the probability of C at X_1, q the probability of C at X_2, and $\sigma = (\sigma_{e_1}, \sigma_{e_2}, \sigma_{e_3}, \sigma_{e_4}) = (p, q, q, p)$. Furthermore, the decision maker's beliefs about the chance player are $\rho(e_1|c) = 0.5$ and $\rho(e_2|c) = 0.5$. It is easy to calculate that the optimal strategy for the decision problem is $(p^\star, q^\star) = (2/3, q_0)$ or $(p^\star, q^\star) = (p_0, 2/3)$ where p_0, q_0 can be arbitrary number in $[0,1]$. Now assume that the decision maker chooses $(p^\star, q^\star) = (2/3, 2/3)$ to implement. In this case, the attitude modified IS-path probability function conditional at X_1 is

$$\Phi(\sigma|X_1,c) = f_{X_1,\sigma}\left(\mu(e_1|X_1,\sigma)((1-p)z_1 + p\lambda(e_1,C|\sigma)\Phi(\sigma|X_2)) + \mu(e_4|X_1,\sigma)((1-p)z_2' + pz_3')\right),$$

where

$$\lambda(e_1,C|\sigma) = \frac{p+1}{p},$$

$$f_{X_1,\sigma}(\cdot) = \sum_{z\in Z_{IS}(X_1)} \frac{1_{\{z\}}(\cdot)\cdot z}{1_{\{z_1\}}(z) + 1_{\{z_2'\}}(z) + 1_{\{z_3'\}}(z) + 1_{Z_{IS}(X_2,\sigma)}(z)}.$$

Here, $Z_{IS}(X_1) = Z_{IS}(X_2) = \{z_1, z_2, z_3, z_1', z_2, z_3'\}$ since we have $X_1 \prec' X_2$ as well as $X_2 \prec' X_1$. Similarly, the attitude modified IS-path probability function conditional at X_2 is

$$\Phi(\sigma|X_2,c) = f_{X_2,\sigma}\left(\mu(e_2|X_2,\sigma)((1-q)z_1' + q\lambda(e_2,C|\sigma)\Phi(\sigma|X_1)) + \mu(e_3|X_2,\sigma)((1-q)z_2 + qz_3)\right),$$

where

$$\lambda(e_2,C|\sigma) = \frac{q+1}{q},$$

$$f_{X_2,\sigma}(\cdot) = \sum_{z\in Z_{IS}(X_1)} \frac{1_{\{z\}}(\cdot)\cdot z}{1_{\{z_1\}}(z) + 1_{\{z_2'\}}(z) + 1_{\{z_3'\}}(z) + 1_{Z_{IS}(X_2,\sigma)}(z)}.$$

Then, apply $\Phi(\sigma|X_1,c)$ and $\Phi(\sigma|X_2,c)$ to the substitution function and we get

$$E^{IS}(p|X_1,c,q^\star) = \frac{1}{1+q^\star}(4p + 4q^\star - 6pq^\star),$$

and

$$E^{IS}(q|X_1,c,p^\star) = \frac{1}{1+p^\star}(4p^\star + 4q - 6p^\star q).$$

Therefore, the decision maker's optimal choice of p is arbitrary based on the fact that $q^\star = 2/3$. He has no incentive to deviate from $p^\star = 2/3$. It is the same when the decision maker at X_2. The ex-ante optimal strategy is time consistent.

There is another way to deal with the cross-branch information sets. In that method, we would like to reconsider the decision problem based on the fact that we are at information set $X_1 \cup X_2$. It will be discussed in a later section.

2.6 IS-time Consistency and One Information Set Deviation Property

We should emphasise that the decision maker cannot convey the information that the strategy has been switched from one node to another node in the same non-singleton information set. If the non-singleton information set does not present absentmindedness, different nodes in the same information set do not belong to one branch of the decision problem tree. It is impossible that the decision maker passes nodes in different paths at one occurrence of execution. Thus, the decision maker cannot control the behavioural rule assigned at other nodes in the same information set. If the information set presents absentmindedness and the decision maker changes strategy at node h in that information set, he cannot notice the switch when he moves to node h', $h \prec h'$ and h' also belongs to that information set. Suppose the decision maker can notice the switch at node h', he must know he has past the node h which is contradicted to the definition of absentmindedness. Absentmindedness indicates that the decision maker cannot even know whether he has past some nodes or not. Therefore, at an information set, the decision maker is only able to control his immediate behaviour. Furthermore, if the decision maker is assumed not to do the reconsideration at following information sets, he should have control over the behaviours assigned at those information sets. In conclusion, a decision maker has control over the immediate behaviour in the current information set X and the behaviours at information sets X' such that $X \prec' X$.

Since the control power of the decision maker has been specified, we should demonstrate all the implicit assumptions in the recursive calculation rule. The decision maker believes the following:

1. Some external process selects information sets where the decision maker reconsiders, one of them is denoted by X. The decision maker does the reconsideration when he knows he is in X and might reconsider somewhere else after he leaves X. The reason why we say it is an external process who chooses the information set is that the decision maker does not know where he will be asked to reconsider during the execution. Thus, the conditional optimality should be satisfied at each information set which is required in the definition of time consistency.

2. Before consideration, the decision maker remembers his initial strategy σ.

3. If the decision maker evaluates a switched behavioural strategy σ' at a node in information set X, he expects the behavioural rule of information set assigned to other nodes is also σ'_X and the conditional expected payoff at other nodes is $E^{IS}u(\sigma'|X)$.

4. From the perspective of the third party, the decision maker reconsiders at each nodes in X along a decision problem tree path. The IS expected payoffs of the information set X is identical at any node or collection of

nodes in X.

5. After reconsidering, the decision maker switches to a different strategy σ', he will remember the new strategy σ' (and then may forget σ).

In the discussion about deviations, the switch of a strategy should be cautiously defined.

Notation 2.6.1 *Denote $[\sigma, X, \sigma']$ the behavioural strategy that the decision maker plays σ' at information set X and every information set X' such that $X \prec' X'$ and no $X' \prec' X$, otherwise plays σ.*

In other words, the decision maker plays σ if he is at an information set X' such that both $X \prec' X'$ and $X' \prec' X$ satisfy. In that case, both σ and σ' are required to be behavioural strategy. It might arise problem when analysing an information set presenting absentmindedness, denoted X_0. Being at a node in X_0, the decision maker might move to another node in the same information set. The expected payoffs at each node is identical to the expected payoffs at the information set. If the decision maker's current optimal behaviour is different from the strategy assumed in the expected payoff function at another node in the same information set, the resulting behavioural rule assigned at X_0 violates the definition of a strategy. However, compared with the modified multiself approach which assumes the decision maker plays the initial strategy at other occurrences, we assume the decision maker will play whatever strategy is implemented now. Thus, the behavioural rule at different nodes in X_0 must be the same.

In comparison, there is another way to define deviations.

Notation 2.6.2 *Define $[\sigma; (\sigma', X)]$ the behavioural strategy that the decision maker plays σ' at information set X and all the information sets X' such that there is a IS-path from X to X' while no IS-path from X' to X, otherwise he plays σ. It is different from $[\sigma, X, \sigma']$.*

For example, as shown in Figure 3.8, if the decision maker plays $[\sigma; (\sigma', \{d_1\})]$, he also plays σ' at information set X_4. Although $X_4 \prec' \{d_1\}$ is not valid, we can see $X_1 \in P_{IS}(X_4)$ and $\{d_1\} \in P'_{IS}(X_1)$, there is an IS-path from $\{d_1\}$ to X_4, and thus the decision maker plays σ' at information set X_4. In comparison, $[\sigma, \{e_1\}, \sigma']$ indicates the decision maker plays σ at information set X_4 since he cannot reach X_4 by implementing any strategy starting from $\{e_1\}$.

Now we define IS-time consistency which follows the inspiration of time consistency under IS expected payoff function.

Definition 2.6.1 *A strategy σ is **IS-time consistent** if*
1) there is a belief $\mu(\sigma)$ weakly consistent with σ;
2) for all information set X which can be reached with positive probability under σ, we have

$$E^{IS}u(\sigma|X) \geq E^{IS}u([\sigma;(\sigma',X)]|X),$$

for any behavioural strategy σ' and its weakly consistent belief $\mu([\sigma;(\sigma',X)])$.

Notice that, in the definition, the strategy to which the decision maker deviates is in the form of $[\sigma;(\sigma',X)]$ rather than a totally different strategy σ'. In decision problems with perfect recall, it is equivalent between deviation to $[\sigma;(\sigma',X)]$ and deviation to σ' since the decision maker's optimal play at one node is independent of his action at any other node and his belief on a non-singleton information set is not affected by his strategy. However, with imperfect recall, the decision maker's belief on a non-singleton information set could be determined by his previous or current behaviours. It is the reason why we need carefully define $[\sigma;(\sigma',X)]$ and use it as the deviating strategy. Another subtlety is that, when evaluating whether to deviate from a strategy, we should find a belief system which is weakly consistent with $[\sigma;(\sigma',X)]$ and use it to calculate the conditional expected payoffs assigned at information set X. It is also because, with absentmindedness, the belief system $\mu(\sigma)$ might not be weakly consistent with $[\sigma;(\sigma',X)]$ at information set X. Thus, the definition of IS-time consistency should include an additional requirement of the existence of a belief weakly consistent of $[\sigma;(\sigma',X)]$.

The same as time consistency, IS-time consistency examines the optimality of a strategy at each information set but not ex-ante. However, if the decision maker moves immediately after the decision problem starts and the initial information set is a singleton, the expected payoff function Eu at the initial information set is identical to ex-ante expected payoff function. In such a scenario, a strategy is ex-ante optimal if it is time consistent. Compared, if the decision maker moves immediately after the execution of the decision problem, no matter how many nodes are included in the initial information set, the expected payoff function of the recursive calculation rule $E^{IS}u$ at the initial information set is the same as ex-ante expected payoff function. Thus, a strategy is ex-ante optimal if it is IS-time consistent. The IS-time consistency indicates there does not exist an alternative strategy that yields higher expected payoffs rather than the initial strategy conditional upon the realisation of any information set. If a behavioural strategy satisfies the criterion, the distinction that the decision maker chooses his behavioural rule between before the play and during execution becomes inconsequential if he implements it and his beliefs are consistent with the strategy.

Proposition 2.6.1 *A strategy is IS-time consistent if and only if it is ex-ante optimal, if the decision problem has no cross-branch information sets.*

Proof of Proposition 2.6.1. See Proposition 2.8.1 and the proof of Proposition 2.8.2.

Proposition 2.6.1 indicates that, with the assumption of rationality (an expected payoff maximiser), there will not be a better alternative strategy at any information set than the initial strategy he chooses before the play. However, an IS-time consistent strategy is not necessarily ex-ante strategy if the decision problem has cross-branch information sets. For example, if we change the payoffs at z_1, z_1', z_2, z_2' in the example of Gilboa Formation (see Figure 3.7) into $u(z_1) = 4, u(z_1') = 4, u(z_2) = 0, u(z_2') = 0$. The strategy to play C at both information sets with probability 1 is IS-time consistent but not ex-ante optimal. We call it revised Gilboa Formation.

We say the decision maker has an initial information set X_0, if all the history h satisfying that there is no history h', $h' \prec h$ and the decision maker moves at h' belongs to the same information set X_0. Then, we could conclude a stronger version of proposition 1 that, in a decision problem where the decision maker has an initial information set, ex-ante optimality is equivalent to IS-time consistency. It is because the probability of reaching X_0 is 1. The IS expected payoff function at the information set X_0 is the same as ex-ante expected payoff function.

Definition 2.6.2 *A strategy σ is* ***time consistent*** *if there is a weakly consistent belief $\mu(\sigma)$ such that for every information set X which is reached with positive probability under σ,*

$$\sum_{h\in X}\mu(h)\sum_{z\in Z}p(z|h,\sigma)u(z) \geq \sum_{h\in X}\mu(h)\sum_{z\in Z}p(z|h,\sigma')u(z),$$

for any behavioural strategy σ'.

Proposition 2.6.2 *For a decision problem with perfect recall, IS-time consistency is equivalent to time consistency.*

Proof of Proposition 2.6.2 We have demonstrated that the recursive calculation rule is identical to the traditional calculation rule with perfect recall, i.e. $E^{IS}u(\sigma|X) = Eu(\sigma|X)$. Thus, a strategy is IS-time consistent if and only if it is time consistent.

Other modification of time consistency was proposed to examine whether a strategy is optimal conditional at each information set. Constrained time consistency in Battigalli (1997) indicates that, at any information set X, a strategy σ which satisfies the property is conditionally optimal over the restricted available strategy set in which any strategy is identical to σ at all information sets except X. The concept explains that the decision maker could notice deviation from the initial strategy only if he reaches an information set which should not be reached if he implements the initial strategy. An ex-ante optimal strategy is constrained time consistent. Thus, ex-ante optimality can deduce conditional optimality over a restricted strategy set. Another concept is gt time consistency in Halpern (1997). A behavioural strategy σ is **gt time consistent** (**game tree time consistent**) if for any information set X which is reached with positive probability and all strategies σ'', we have

$$\sum_{h\in\hat{X}}\mu(h|\hat{X},\sigma)E(\sigma|h) \geq \sum_{h\in\hat{X}}\mu(h|\hat{X},\sigma'')E(\sigma'|h).$$

It indicates, at any information set X, the decision maker reconsiders only at the upper frontier. An ex-ante optimal strategy is gt time consistent. Furthermore, an ex-ante optimality is equivalent to gt time consistency if the decision problem presents partial recall. Define $exp''(h)$ for a history h consists of information sets that the decision maker passes until reaching h without consecutive repetitions. A decision problem is said to have partial recall, if for any information set X and two nodes $h,h'\in X$, we have $exp''(h) = exp''(h')$. In other words, partial recall excludes the situation that the decision maker forgets some previous acquired information. The information could be the chance player's behaviours or previous behaviours of the decision maker himself. With partial recall, IS-time consistency is equivalent to gt time consistency. Besides, IS-time consistency elucidates an ex-ante optimal strategy is still conditionally optimal even we assume the decision maker could deviate to any available strategy and extend the restriction on decision problems from partial recall to any assumption.

Rather than discussion over one shot deviation, we propose one information set deviation in this paper. The modified multiself consistency in PR explains a strategy satisfies the property if the decision maker cannot achieve higher expected payoffs by changing his immediate action. The environment for modified multiself approach is that the external process picks a unique node in an information set, and asks the decision maker to reconsider. The decision maker will not reconsider anywhere else. The decision maker cannot reconsider at only one node from an information set with absentmindedness. Suppose the decision maker knows he will reconsider at information set X. When he is at a node in X, he will reconsider at each node along a path since he does not know even whether he has passed a node in X or not. Reevaluation at a node in an information set with absentmindedness can be realised if the decision maker only considers such an issue when he is stopped by a third party during execution. However, we should know PR excludes the possibility of absentmindedness. There is no controversy in their framework that one-shot deviation can be realised by the decision maker. Furthermore, the recursive calculation rule shows the decision maker's attitude on information sets. He regards an information set as a whole unit. When he evaluates a change in his immediate behaviour, he expects the same change should occurs also at other nodes in the same information set. The recursive calculation rule evaluates the change of behaviour on the whole information set instead of a particular node. It is the reason why we analyse one information set deviation in this paper.

Besides our framework, the concept "ms (multiself) time consistency" in Halpern (1997) also evaluates one information set deviation. A behavioural strategy σ is **ms time consistent** if for any information sets X and all strategies σ' which assign identical behavioural rule at all information sets except X, i.e. $\sigma'(X') \neq \sigma(X')$ if $X' \neq X$, we have $Eu(\sigma|\emptyset) \geq Eu(\sigma'|\emptyset)$. Ms time consistency is equivalent to modified multiself consistency if no absentmindedness presents in a finite decision problem. It is actually the evaluation of one information set deviation under the gt time consistency. It is easily to see that

$$\sum_{h\in\hat{X}} \mu(h|\hat{X},\sigma)E(\sigma|h) = \frac{p(h|\emptyset,\sigma)E(\sigma|h)}{\sum_{h\in\hat{X}} p(h|\emptyset,\sigma)} \quad = \frac{\sum_{z\in Z(\hat{X})} p(z|\emptyset,\sigma)u(z)}{p(X|\emptyset,\sigma)}$$

where $Z(\hat{X})$ is the set of terminal nodes which can be reached by a completely mixed behavioural strategy from the nodes in the upper frontier of information set X. The term $P(X|\emptyset,\sigma)$ is not changed by the alternation from σ to σ', between which the alternation of strategy occurs at information set X. Then, the inequation that expresses gt time consistency is equivalent to the following inequation.

$$\sum_{z\in Z(X)} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z(X)} p(z|\emptyset,\sigma'),$$

$$\sum_{z\in Z} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z} p(z|\emptyset,\sigma').$$

Thus, we have $Eu(\sigma|\emptyset) \geq Eu(\sigma'|\emptyset)$. If a strategy satisfies ms time consistency, it must satisfy gs time consistency. In other words, ms time consistency describes, being in the information set X, the decision maker is stopped at the upper frontier and asked to reconsider his behavioural rule on the current information set and will not be stopped again anywhere else. The decision maker can control his behaviours at the whole information set. Conditional being at a node in $X/\hat{X}$, he is informed of the latest strategy he switches at the upper frontier of X. However, not only reconsideration solely at the upper frontier but also being informed of the latest switched strategy should be done by an external observer. The decision maker is not able to make it himself. It is the reason why we establish the framework based on the recursive calculation rule.

Definition 2.6.3 *A strategy σ satisfies* ***one information set deviation*** *property if*
1) there is a belief $\mu(\sigma)$ weakly consistent with σ;
2) for all information set X which can be reached with positive probability under σ, and any behavioural rule σ'_X on the information set X, there is a belief $\mu((\sigma'_X,\sigma_{-X}))$ weakly consistent with (σ'_X,σ_{-X}), we have

$$E^{IS}u(\sigma|X) \geq E^{IS}u((\sigma'_X,\sigma_{-X})|X).$$

We have known that, in a finite decision problem with imperfect recall, one-shot deviation is not equivalent to time consistency. The equivalence is still not valid even the deviation is extended from the decision maker's immediate behaviour to the whole information set.

Proposition 2.6.3 *A strategy σ which is IS-time consistent satisfies one information set deviation property, but not vice versa.*

Proof of Proposition 2.6.3. $(\sigma''_X, \sigma_{-X})$ is a special case of $[\sigma; (\sigma', X)]$, in which $\sigma' = (\sigma''_X, \sigma_X)$. Thus, one information set deviation property is a special case of IS-time consistency.

Corollary 2.6.1 *In a finite decision problem, strategy σ satisfies one information set deviation property if it is ex-ante optimal, but not vice versa.*

Now we take an example that a strategy satisfies one information set deviation property but is not IS-time consistent. In example 1 (see Figure 3.4), the strategy $(s_{e_1}, s_{e_2}, s_{e_3}) = (L, l, l)$ satisfies one information set deviation property. Based on the knowledge that the decision maker played L at e_1, the expected payoff at information set K is q, if the probability of playing l at K is q. Then, the decision maker should choose to play l with probability 1. At e_1, the expected payoff function is $pq + 2(1-p)(1-q)$, if the probability of playing L at e_1 is p. Based on the knowledge that he will play l at K. The expected payoff becomes p. Then, it is optimal that he should play L at e_1. However, the only IS-time consistent strategy is $(s_{e_1}, s_{e_2}, s_{e}3) = (R, r, r)$ in example 1. Thus, the example shows one information set deviation property cannot deduce IS-time consistency.

Proposition 2.6.4 *One information set deviation property is equivalent to ms time consistency, if there do not exist cross-branch information sets in the finite decision problem.*

Proof of Proposition 2.6.4. See the Appendix.

Both one information set property and ms time consistency describe the same deviation at each node in the same information set. However, they interpret it in different way. One information set property holds that the decision maker can only control the immediate action but expect he would deviate to the same behavioural rule at other nodes while ms time consistency demonstrates the decision maker only reconsiders at the upper frontier and he knows the updated strategy after the reconsideration and follows it at every node in that information set. However, as we stated above, to only reconsider at the upper frontier cannot be achieved by the decision maker himself. There must be an external manipulator who stops the decision maker to reconsider, otherwise the decision maker would follow the latest strategy without hesitation. There exists another question. How does the decision maker know the latest strategy updated at the upper frontier if he even does not remember he has experienced some node in the upper frontier? Thus, it seems that the interpretation of one information set property is more reasonable. In fact, there is a potential problem in the framework of recursive calculation rule. When there are ties in the set of conditional optimal strategies, it cannot be guaranteed that the decision maker expects the same strategy at different node in the same information set.

Back to the distinction between agent form and multiself approach of a decision problem when absentmindedness presents, the ms time consistency and one information set property represents the perspective of agent form while modified multiself consistency shows the analysis of multiself approach.

Corollary 2.6.2 *In a finite decision problem without absentmindedness, one information set deviation property is equivalent to modified multiself consistency.*

Proof of Corollary 2.6.2. One information set property is equivalent to ms time consistency. Without absentmindedness, ms time consistency is equivalent to modified multiself consistency. Thus, one information set property is equivalent to modifies multiself consistency.

Corollary 2.6.3 *In a finite decision problem with perfect recall, one information set deviation property is equivalent to one-shot deviation principle.*

2.7 Strongly IS-time Consistency

With assumptions of full rationality and implementation of a pure strategy, the decision maker might acquire some information regarding his current position in a non-singleton information set. Therefore, if imperfect recall appears, we should relax the assumption of rationality. The decision maker might make mistake either when he chooses a strategy, and then cause the chosen strategy not conditionally optimal; or when he implements the chosen strategy, for example, the decision maker is trembling-hand. When potential mistake is allowed, the chosen strategy has no direct implication on the decision maker's position in the decision problem tree. Like in example 2 (see Figure 3.5), knowing the latest strategy which requires to implement L at information set X, the decision maker should know he is current at e_3 if he learns he had reconsidered at e_1 or e_2 and was rational when do the reconsideration. However, if potential mistake is allowed, the decision maker cannot confirm he is at e_3 although he knows the latest strategy. In this case, the fact that the decision maker is assumed to remember the latest switched strategy could hardly convey more information than the strategy itself. Another reason to relax assumptions is that the decision maker maker should have best response even if he finds himself at some unexpected information sets.

Definition 2.7.1 *We say a belief system μ is* ***strongly consistent*** *with a behavioural strategy σ if there exists a sequence of belief systems μ_k which are strictly positive weakly consistent with completely mixed behavioural strategies σ_k respectively, such that $lim_{k\to\infty}(\mu_k,\sigma_k) = (\mu,\sigma)$.*

Definition 2.7.2 *A strategy σ is* ***strongly IS-time consistent*** *if*
1) there is a belief $\mu(\sigma)$ strongly consistent σ, such that,
2) for all information sets X, we have

$$E^{IS}u(\sigma|X) \geq E^{IS}u([\sigma;(\sigma',X)]|X),$$

for any behavioural strategy σ' and its weakly consistent belief $\mu([\sigma;(\sigma',X)])$.

Proposition 2.7.1 *In a finite decision problem, for any ex-ante optimal strategy, there is an outcome-equivalent behavioural strategy satisfying strongly IS-time consistency.*

Proof of Proposition 2.7.1. See Proposition 2.8.1 and the proof of Proposition 2.8.3.

The assessment (μ, σ) is composed by an ex-ante optimal strategy σ and its strongly consistent belief system $\mu(\sigma)$ meets the requirement of sequential rationality if the decision problem has no cross-branch information sets. An assessment (μ, σ) is sequentially rational if the strategy σ maximises the conditional expected payoffs at every information set X given the belief system μ. However, our sequential rationality is established on the expected payoff function of $E^{IS}u$ instead of Eu. Furthermore, if a decision problem has cross-branch information sets, the strongly IS-time consistency cannot guarantee a strategy is conditionally optimal over the whole behavioural strategy set. When a decision maker is at one of cross-branch information sets, the available behavioural strategy set excludes those which changes his behaviours at the other cross-branch information sets. Thus, it is necessary to define the decision maker's deviations at multiple information sets.

2.8 Reconsideration at Multiple Information Sets

As we have mentioned in previous section, another way to deal with cross-branch information sets is to allow simultaneous changes at multiple information sets. For example, in the example of Gilboa's formation, the decision maker could also evaluate the decision problem at $X_1 \cup X_2$. Then, the decision maker is able to control the behaviours at both information sets. We only discuss the situation when $p, q > 0$ here. The attitude modified IS-path probability function is

$$\begin{aligned}
\Phi(\sigma|X_1 \cup X_2, c) &= \mu(e_1|X_1 \cup X_2, \sigma, c)[(1-p)z_1 + p\Phi(\sigma|X_1 \cup X_2, c)] \\
&+ \mu(e_2|X_1 \cup X_2, \sigma, c)[(1-q)z_1' + q\Phi(\sigma|X_1 \cup X_2, c)] \\
&+ \mu(e_3|X_1 \cup X_2, \sigma, c)[(1-q)z_2 + qz_3] \\
&+ \mu(e_4|X_1 \cup X_2, \sigma, c)[(1-p)z_2' + pz_3'] \\
&= \frac{1}{2+p+q}[(1-p)z_1 + p\Phi(\sigma|X_1 \cup X_2, c)] + \frac{1}{2+p+q}[(1-q)z_1' + q\Phi(\sigma|X_1 \cup X_2, c)] \\
&+ \frac{p}{2+p+q}[(1-q)z_2 + qz_3] + \frac{q}{2+p+q}[(1-p)z_2' + pz_3'].
\end{aligned}$$

if the decision maker's strategy σ indicates to choose C with probability p at X_1 and choose C with probability q at X_2 Then, the function is

$$\begin{aligned}\Phi(\sigma|X_1 \cup X_2, c) &= \frac{1}{p(X_1 \cup X_2|\emptyset, \sigma, c)} \sum_{z \in Z_{IS}(X_1 \cup X_2)} p(z|\emptyset, \sigma) z \\ &= \frac{1}{2}[(1-p)z_1 + (1-q)z_1' + p(1-q)z_2 + pqz_3 + q(1-p)z_2' + pqz_3'].\end{aligned}$$

We apply it to the substitution function and get

$$E^{IS}u(\sigma|X_1 \cup X_2, c) = \frac{1}{p(X_1 \cup X_2|\emptyset, \sigma, c)} \sum_{z \in Z_{IS}(X_1 \cup X_2)} p(z|\emptyset, \sigma)u(z) = 2p + 2q - 3pq.$$

It is easy to know there is no other strategy which reaches higher expected payoffs than the ex-ante optimal strategy. Thus, the ex-ante optimal strategy is time consistent.

Another example which we could assume the decision maker reconsiders at multiple information sets is "Forget previous acquired information" (see Figure 3.10). The decision maker is assumed to reconsider at node e_1 and e_2 simultaneously. Then, the attitude modified IS-path probability function is

$$\begin{aligned}\Phi(\sigma|\{e_1\} \cup \{e_2\}) &= \mu(e_1|\{e_1\} \cup \{e_2\}, \sigma, c)[(1-p_1)z_1 + p_1\Phi(\sigma|X, c)] \\ &\quad + \mu(e_2|\{e_1\} \cup \{e_2\}, \sigma, c)[(1-p_2)z_2 + p_2\Phi(\sigma|X, c)] \\ &= \frac{1}{2}[(1-p_1)z_1 + (1-p_2)z_2 + (p_1+p_2)\Phi(\sigma|X, c)]\end{aligned}$$

where

$$\Phi(\sigma|X, c) = \frac{2}{p_0(p_1+p_2)} \sum_{i=\{3,4,5,6\}} p(z_i|\emptyset, \sigma, c)z_i.$$

Then, the function $\Phi(\sigma|\{e_1\} \cup \{e_2\})$ is

$$\begin{aligned}\Phi(\sigma|\{e_1\} \cup \{e_2\}) &= \frac{1}{p(\{e_1\} \cup \{e_2\}|\emptyset, \sigma, c)} \sum_{z \in Z_{IS}(\{e_1\} \cup \{e_2\})} p(z|\emptyset, \sigma, c)z \\ &= \frac{1}{p_0} \cdot \frac{1}{2} p_0[(1-p_1)z_1 + (1-p_2)z_2 + p_1pz_3 + p_1(1-p)z_4 + p_2pz_5 + p_2(1-p)z_6].\end{aligned}$$

Notice that. the decision maker's strategy σ is $(\sigma_{e_0}(G), \sigma_{e_1}(D), \sigma_{e_2}(D), \sigma_{e_3}(L), \sigma_{e_4}(L)) = (p_0, p_1, p_2, p, p)$.

Then, apply it to the substitution function, we get

$$E^{IS}u(\sigma|X,c) = \frac{1}{2}[2(1-p_1)+2(1-p_2)+5p_1p+6p_2(1-p)].$$

It is easy to know there is no other strategy which reaches higher expected payoffs than the ex-ante optimal strategy. Thus, the ex-ante optimal strategy is time consistent.

Now we develop the general framework of reconsideration at multiple information sets simultaneously. Denote $\mathscr{X}_1 \subseteq \mathscr{X}$ a subset of $\mathscr{X}$.

Notation 2.8.1 *For two sets of information sets $\mathscr{X}_1$ and $\mathscr{X}_2$, we say $\mathscr{X}_1 \prec' \mathscr{X}_2$ if there is no information set X such that $X \in \mathscr{X}_1 \cap \mathscr{X}_2$ and $\forall X_2 \in \mathscr{X}_2, \exists X_1 \in \mathscr{X}_1$, such that $X_1 \prec' X_2$.*

The decision maker reconsiders at information sets in $\mathscr{X}_1$. A generalised belief system μ assigns a probability $\mu(h|\mathscr{X}_1)$ to a history h in one of the information sets in $\mathscr{X}_1$. It can be interpreted as the probability of reaching history h conditional on reaching the union of information sets, $\mathscr{X}_1$. Denote $\widehat{\mathscr{X}_1}$ the upper frontier of $\mathscr{X}_1$, then the probability of reaching $\mathscr{X}_1$ is

$$p(\mathscr{X}_1|\emptyset,\sigma) = \sum_{h\in\widehat{\mathscr{X}_1}} p(h|\emptyset,\sigma),$$

if the decision maker implements the behavioural strategy σ.

Result 2.8.1 *Assume that the decision maker implements the behavioural strategy σ, and reconsiders at multiple information sets in the non-empty subset $\mathscr{X}_1 \subseteq \mathscr{X}$ simultaneously with the belief system μ. Then, the attitude modifies IS-path probability function is*

$$\Phi(\sigma|\mathscr{X}_1) = \frac{1}{P(\mathscr{X}_1|\emptyset,\sigma)} \sum_{z\in Z_{IS}(\mathscr{X}_1,\sigma)} p(z|\emptyset,\sigma)z,$$

and its corresponding IS expected payoff function is

$$E^{IS}u(\sigma|\mathscr{X}_1) = \frac{1}{P(\mathscr{X}_1|\emptyset,\sigma)} \sum_{z \in Z_{IS}(\mathscr{X}_1,\sigma)} p(z|\emptyset,\sigma)u(z).$$

There is a feasible IS-path from $\mathscr{X}$ to the set of terminal node z, $\{z\}$, if there exists at least one feasible IS-path from an information set $X \in \mathscr{X}$ to $\{z\}$.

Proof of Result 2.8.1. Similar to the proof of Result 2.5.2.

So far, we have described how to model reconsideration at multiple information sets simultaneously in a decision problem with imperfect recall in the relation $\prec'$. An ex-ante optimal strategy is time consistent even the decision maker reconsiders at multiple information sets simultaneously since the resulting expected payoff function is consistent with the IS expected payoff function at an information set.

We know that the altered strategy $[\sigma, X, \sigma']$ or $[\sigma; (\sigma', X)]$ from a strategy σ requires the decision maker plays σ' at information set X' if $X \prec' X'$ and no $X' \prec X$. As we have analysed, at one of cross-branch information sets, the decision maker has no control over the behaviours at the others. It is hard to examine whether a strategy is sequentially rational based on the fact that the decision maker can change the behaviours at all cross-branch information sets. Thus, allowing simultaneous changes at multiple information sets allow us to examine a stronger requirement of time consistency.

Definition 2.8.1 *A strategy σ is* ***modified (strongly) IS-time consistent*** *if*
1) there is a belief $\mu(\sigma)$ weakly (strongly) consistent σ,such that,
2) for all sets of information sets $\mathscr{X}_1$ which can be reached with positive probability under σ (for all sets of information sets $\mathscr{X}_1$), we have

$$E^{IS}u(\sigma|\mathscr{X}_1) \geq E^{IS}u([\sigma; (\sigma', \mathscr{X}_1)]|\mathscr{X}_1),$$

for any behavioural strategy σ', the behavioural strategy $[\sigma;(\sigma',\mathscr{X}_1)]$ and its weakly consistent belief $\mu([\sigma;(\sigma',\mathscr{X}_1)])$

Proposition 2.8.1 *A strategy is (strongly) IS-time consistent if it is modified (strongly) IS-time consistent.*

Proof of Proposition 2.8.1. When the set of information sets $\mathscr{X}_1$ in the definition of modified (strongly) IS-time consistency contains only one information set X, the modified (strongly) IS-time consistency is the same as (strongly) IS-time consistency.

Proposition 2.8.2 *A strategy is modified IS-time consistent if and only if it is ex-ante optimal.*

Proof of Proposition 2.8.2. See the Appendix.

Proposition 2.8.3 *For each ex-ante optimal strategy, there is an outcome-equivalent behavioural strategy satisfying modified strongly IS-time consistency.*

Proof of Proposition 2.8.3. See the Appendix.

Now we have defined the decision maker's deviations at multiple information sets. Following the discussion of sequential rationality in previous section, a strategy which is modified strongly IS-time consistent satisfies sequential rationality.

2.9 Conclusions and Discussion

Due to the time inconsistency of an ex-ante strategy in decision problems with imperfect recall, we start to focus on the execution process of a decision problem. If the decision maker is stopped in the middle of the play and allowed to change his strategy, he must evaluate the conditional expected payoffs of available strategies. Furthermore, if the ex-ante strategy is different from the conditional optimal one at the information set where he is stopped, the decision maker can choose to deviate to a different strategy or believe his initial choice and do not change. More important, whether there exists an ex-ante strategy which is also optimal conditional on the fact that the decision maker is at every information set. Thus, how the decision maker evaluates the expected payoffs becomes especially significant.

In this paper, we propose the sophisticated recursive calculation rule to calculate the expected payoffs conditional at an information set or multiple information sets. Different from the traditional calculation rule, we treat an information set as a whole unit, and the decision maker moves one step to an information set instead of a history. Additionally, the decision maker estimates the conditional probability to a terminal node through a path of information sets. At a non-singleton information set, the decision maker forms beliefs of the probability being at each history in that information set. Based at each history, the decision maker's different behaviours lead him to different information sets. The information set that the decision maker reaches after one step move could be the same information set before the move. In that case, the information set presents absentmindedness. The product of the probability of a particular action and the attitude modified expected payoffs of the information set which the decision maker reaches after taking that action is the expected payoffs of that particular action at a history. The sum of the expected payoffs over all available actions at a history is the expected payoffs of that history. Then, the weighted sum of expected payoffs at each history in a non-singleton information set, in which the weight is the decision maker's belief, is the expected payoffs at that information set. Therefore, the probability to reach a terminal node is estimated as the modified probability from the current information set to other information sets until the terminal node, where the two adjacent information sets in this path satisfy a history in an information set is the subhistory of a history in the other information set. The decision maker knows how he evaluates the expected payoffs if he is stopped again in the later execution of the decision problem and takes it into consideration when he reconsiders his strategy currently. We say those decision makers are sophisticated. In our framework, a terminal node might be evaluated more than once through different IS-paths, then we assign equal probability to each IS-path since the decision maker regards an information set as a complete unit and does not know which one can really reach that terminal node. Therefore, the IS expected payoff function is defined.

Then, in decision problems without cross-branch information sets, we have defined IS-time consistent. If a strategy satisfies IS-time consistency, it is optimal conditional at all information sets that could be reached by that strategy. Of course, the conditional optimality is evaluated in terms of IS expected payoff function. Additionally, strongly IS-time consistency is also defined. There exists an ex-ante strategy which is strongly IS-time consistent indicates a strategy which satisfies sequential rationality exists. However, if a decision problem has cross-branch information sets, (strongly) IS-time consistency is not enough to show the conditional optimality of a strategy. Thus, we define a decision maker's deviations at multiple information sets and modified (strongly) IS-time consistency. In the meanwhile, we define one information set deviation property which shows the contradicted attitude of the deviations at an information set with absentmindedness with modified multiself consistency. Modified multiself consistency assumes that the decision maker will follow the initial strategy at the future occurrence of the current information set while one information set deviation property assumes the decision maker would take whatever actions he adopts now at his future arrival of the current information set.

This paper discusses the ambiguous interpretation of the decision problems with imperfect recall, especially during the process of execution and defines a new method to calculate conditional expected payoffs. In fact, most works on imperfect recall advocate more precise knowledge of histories in a non-singleton information set. To analyse decision problems in terms of information sets shows the reverse trend. However, our analysis is based on the fact that the decision maker knows how he would evaluate at his future reconsideration. It describes how a sophisticated decision maker thinks when he is stopped at an information set.

So far, the discussions are restricted to decision problems. It is also interesting to incorporate the framework into games. In games, it relates to not only a player's own mistakes but also potential mistakes by other players. Additionally, the imperfect recall we discuss in this paper presents quite a simple situation. The decision maker cannot figure out the accurate history he is currently located in an information set. However, there could be another type of imperfect recall which assigns asymmetric information to different nodes. Like in the absentminded driver paradox, the decision maker might be aware of his current position when he is at the first intersection but cannot differentiate two intersections when he moves to the intersection e_2 (Geanakoplos 1989). Of course, there are many other more complicated types of imperfect recall and it is hard to just model them.

2.10 Appendix

Proof of Result 2.5.2.

In this proof, the decision problems discussed exclude the chance player to simplify the notation.

Situation 1. The decision problem has no cross-branch information sets.

Step 1. There exists at least one information set X_0 which satisfies $\forall h'$ such that $h \prec h'$, $h \in X_0$, either $h' \in X_0$ or $h' \in Z$. For any information set which satisfies the above conditions, we have

$$f_{X,\sigma}(\cdot) = \sum_{z \in Z_{IS}(X,\sigma)} 1_{\{z\}}(\cdot) \cdot z,$$

for any behavioural strategy σ.

The attitude modified IS-path probability function at X is

$$\begin{aligned}\Phi(\sigma|X) &= \sum_{a\in A(X)} \sigma_X(a) \sum_{h\in X} \mu(h|X,\sigma)\cdot\lambda(h,a|\sigma)\cdot\Phi(\sigma|Y_{(h,a)}) \\ &= \sum_{a\in A(X)} \sigma_X(a) \sum_{h\in X} \mu(h|X,\sigma)\cdot\Phi(\sigma|Y_{(h,a)}) \\ &= \frac{\sum_{h\in X/\hat{X}} p(h|\emptyset,\sigma)}{\sum_{h\in X} p(h|\emptyset,\sigma)}\Phi(\sigma|X) + \sum_{z\in Z_{IS}(X,\sigma)} \frac{p(z|\emptyset,\sigma)}{\sum_{h\in X} p(h|\emptyset,\sigma)}\cdot z\end{aligned}$$

It is easy to know that

$$\frac{\sum_{h\in\hat{X}} p(h|\emptyset,\sigma)}{\sum_{h\in X} p(h|\emptyset,\sigma)}\Phi(\sigma|X) = \sum_{z\in Z_{IS}(X,\sigma)} \frac{p(z|\emptyset,\sigma)}{\sum_{h\in X} p(h|\emptyset,\sigma)}\cdot z,$$

$$\begin{aligned}\Phi(\sigma|X) &= \frac{1}{\sum_{h\in\hat{X}} p(h|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(h|\emptyset,\sigma)\cdot z \\ &= \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(h|\emptyset,\sigma)\cdot z\end{aligned}$$

Apply it to the substitution function, we get

$$E^{IS}u(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(h|\emptyset,\sigma)\cdot u(z).$$

Step 2. By complete induction in mathematics, for an information set X, if for any $h\in X$ and $a\in A(X)$, for those $Y_{(h,a)}\neq X$, we have

$$\Phi(\sigma|X_{(h,a)}) = \frac{1}{p(X_{(h,a)}|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X_{(h,a)},\sigma)} p(z|\emptyset,\sigma)\cdot z$$

if $Y_{(h,a)} = X_{(h,a)}$,

$$\Phi(\sigma|Y_{(h,a)}) = Y_{(h,a)} = z$$

if $Y_{(h,a)} = z$.

Assume that σ is a completely behavioural strategy, then for any $h \in X$, $p_{X_h}(X_{(h,a)})|\emptyset,\sigma) \neq 0$, then

$$\Phi(\sigma|X) = f_{X,\sigma}\left(\sum_{a\in A(X)} \sigma_X(a) \sum_{h\in X} \mu(h|X,\sigma)\cdot\lambda(h,a|\sigma)\cdot\Phi(\sigma|Y_{(h,a)})\right)$$

$$= f_{X,\sigma}\Big(\frac{1}{\sum_{h'\in X} p(h'|\emptyset,\sigma)}\Big[\sum_{h\in X/\hat{X}} p(h|\emptyset,\sigma)\Phi(\sigma|X)$$

$$+ \sum_{\substack{h\in X, a\in A(X), (h,a)\notin X \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} p((h,a)|h,\sigma)p(h|\emptyset,\sigma)\frac{p(Y_{(h,a)}|\emptyset,\sigma)}{p_X(Y_{(h,a)}|\emptyset,\sigma)}\Phi(\sigma|Y_{(h,a)})\Big]\Big)$$

$$\sum_{h'\in\hat{X}} p(h'|\emptyset,\sigma)\Phi(\sigma|X) = f_{X,\sigma}\left(\sum_{\substack{h\in X, a\in A(X), (h,a)\notin X \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} \frac{p((h,a)|\emptyset,\sigma)p(Y_{(h,a)}|\emptyset,\sigma)}{p_X(Y_{(h,a)}|\emptyset,\sigma)} \frac{1}{p(Y_{(h,a)}|\emptyset,\sigma)} \sum_{z\in Z_{IS}(Y_{(h,a)},\sigma)} p(z|\emptyset,\sigma)\cdot z\right)$$

$$= f_{X,\sigma}\left(\sum_{\substack{h\in X, a\in A(X), (h,a)\notin X \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} \frac{p((h,a)|\emptyset,\sigma)}{p_X(Y_{(h,a)}|\emptyset,\sigma)} \sum_{z\in Z_{IS}(Y_{(h,a)},\sigma)} p(z|\emptyset,\sigma)\cdot z\right)$$

$$= f_{X,\sigma}\left(\sum_{\substack{Y_{(h,a)}\neq X \text{ where} \\ h\in X, a\in A(X) \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} \sum_{z\in Z_{IS}(Y_{(h,a)},\sigma)} p(z|\emptyset,\sigma)\cdot z\right)$$

$$= \sum_{z\in Z_{IS}(X,\sigma)} \frac{z}{\sum_{\substack{X\in P'_{IS}(Y), Y\in\mathscr{Y}, \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} 1_{Z_{IS}(Y,\sigma)}(z)} \sum_{\substack{Z_{IS}(Y_{(h,a)}) \text{ s.t.} \\ z\in Z_{IS}(Y_{(h,a)}), Y_{(h,a)}\neq X \\ p_X(Y_{(h,a)}|\emptyset,\sigma)>0}} p(z|\emptyset,\sigma)\cdot z$$

$$= \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot z$$

$$\Phi(\sigma|X) = \frac{1}{\sum_{h'\in\hat{X}} p(h'|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot z$$

$$= \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot z$$

Then, apply it to the substitution function and we get

$$E^{IS}u(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot u(z).$$

Thus, by complete induction, we conclude

$$\Phi(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot z,$$

and

$$E^{IS}u(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot u(z).$$

Situation 2. The decision problem has cross-branch information sets. In this case, it is difficult to calculate Φ directly. We would like to verify the expression

$$\Phi(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)\cdot z$$

instead of doing the calculation. From the calculation in Situation 1., the expression satisfies

$$\Phi(\sigma|X) = f_{X,\sigma}\left(\sum_{a\in A(X)} \sigma_X(a) \sum_{h\in X} \mu(h|X,\sigma)\cdot\lambda(h,a|\sigma)\cdot\Phi(\sigma|Y_{(h,a)})\right).$$

Thus, the expression is the correct attitude modified IS-path probability function, even there are cross-branch information sets in a decision problem.

Proof of Proposition 2.6.4.

If a strategy σ and its weakly consistent belief system $\mu(\sigma)$ satisfies one information set deviation property, we have

$$E^{IS}u(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)u(z) \geq \frac{1}{p(X|\emptyset,\sigma')} \sum_{z\in Z_{IS}(X,\sigma')} p(z|\emptyset,\sigma')u(z) = E^{IS}u(\sigma'|X).$$

The strategy σ' agrees with σ at all information sets except X and could be arbitrary behavioural rule σ'_X on X. Require that σ'_X is completely mixed, i.e., $\forall a \in A(X), \sigma'_X(a) > 0$. Denote $\cup_{X \in P'_{IS}(X')} Z_{IS}(X', \sigma)$ ($\cup Z_{IS}(X', \sigma)$ for simplicity in notation) the union of the sets of terminal node z satisfying there is a feasible IS-path from the information set X', where X' is one of the immediate successor information set of X, to the set containing the terminal node z. We have $\cup Z_{IS}(X', \sigma) = \cup Z_{IS}(X', \sigma')$ since, for any X', X and X' are not cross-branch information sets, i.e., for any $h' \in X'$, there is no $h \in X$ such that $h' \prec h$. We have $Z_{IS}(X, \sigma') = \cup Z_{IS}(X', \sigma')$, if $\sigma'_X(\sigma')$ is completely mixed.

Then, we have

$$\sum_{z \in Z_{IS}(X,\sigma)} p(z|\emptyset, \sigma)u(z) \geq \sum_{z \in Z_{IS}(X,\sigma')} p(z|\emptyset, \sigma')u(z) = \sum_{z \in \cup Z_{IS}(X',\sigma')} p(z|\emptyset, \sigma')u(z).$$

Thus,

$$\sum_{z \in \cup Z_{IS}(X',\sigma)} p(z|\emptyset, \sigma')u(z) > \sum_{z \in \cup Z_{IS}(X',\sigma')} p(z|\emptyset, \sigma')u(z)$$

since $Z_{IS}(X, \sigma) \subseteq \cup Z_{IS}(X', \sigma)$. For any $z \in Z \setminus \cup Z_{IS}(X', \sigma) = Z \setminus \cup Z_{IS}(X', \sigma')$, $p(z|h, \sigma') = 0, \forall h \in X, \sigma'$, then

$$\sum_{z \in Z \setminus \cup Z_{IS}(X',\sigma)} p(z|\emptyset, \sigma)u(z) = \sum_{z \in Z \setminus \cup Z_{IS}(X',\sigma')} p(z|\emptyset, \sigma')u(z).$$

In the meanwhile, we have

$$\begin{aligned}\frac{1}{p(X|\emptyset, \sigma)} &= \frac{1}{\sum_{h \in \hat{X}} p(h|\emptyset, \sigma)} \\ &= \frac{1}{\sum_{h \in \hat{X}} p(h|\emptyset, \sigma')} = \frac{1}{p(X|\emptyset, \sigma')}\end{aligned}$$

thus

$$Eu(\sigma|\emptyset) = \sum_{z \in Z} p(z|\emptyset, \sigma)u(z) \geq \sum_{z \in Z} p(z|\emptyset, \sigma')u(z) \geq Eu(\sigma'|\emptyset).$$

Thus, the strategy σ is ms time consistent.

If strategy σ is ms time consistent if for all information sets X and all strategy σ' which is identical to σ at all information sets except X, we have

$$Eu(\sigma|\emptyset) = \sum_{z\in Z} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z} p(z|\emptyset,\sigma')u(z) = Eu(\sigma'|\emptyset).$$

For any σ', and a very small ε, construct $\sigma_n'' = (1-\frac{\varepsilon}{n})\sigma' + \frac{\varepsilon}{n}\sigma, n \in \mathcal{N}$. Then,

$$Eu(\sigma|\emptyset) = \sum_{z\in Z} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z} p(z|\emptyset,\sigma_n'')u(z) = Eu(\sigma_n''|\emptyset), \forall n \in \mathcal{N}.$$

Then, $Z_{IS}(X,\sigma_n'') = Z_{IS}(X,\sigma) \cup Z_{IS}(X,\sigma'), \forall n \in \mathcal{N}$. If $z \in Z \setminus Z_{IS}(X,\sigma_n'')$, $p(z|h,\sigma) = 0, \forall \sigma'$ and $h \in X$, thus for any n

$$\sum_{z\in Z\setminus Z_{IS}(X,\sigma_n'')} p(z|\emptyset,\sigma)u(z) = \sum_{z\in Z\setminus Z_{IS}(X,\sigma_n'')} p(z|\emptyset,\sigma_n'')u(z).$$

Then,

$$\sum_{z\in Z_{IS}(X,\sigma_n'')} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z_{IS}(X,\sigma_n'')} p(z|\emptyset,\sigma_n'')u(z),$$

$$\sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z_{IS}(X,\sigma_n'')} p(z|\emptyset,\sigma_n'')u(z).$$

Additionally,

$$\begin{aligned}\frac{1}{p(X|\emptyset,\sigma)} &= \frac{1}{\sum_{h\in\hat{X}} p(h|\emptyset,\sigma)} \\ &= \frac{1}{\sum_{h\in\hat{X}} p(h|\emptyset,\sigma_n'')} = \frac{1}{p(X|\emptyset,\sigma_n'')}\end{aligned}$$

It is easy to see that when $n \to \infty$, $\sigma_n'' \to \sigma'$. Thus,

$$E^{IS}u(\sigma|X) = \frac{1}{p(X|\emptyset,\sigma)} \sum_{z\in Z_{IS}(X,\sigma)} p(z|\emptyset,\sigma)u(z) \geq \frac{1}{p(X|\emptyset,\sigma')} \sum_{z\in Z_{IS}(X,\sigma')} p(z|\emptyset,\sigma')u(z) = E^{IS}u(\sigma'|X).$$

The strategy σ satisfies one information set deviation property.

Proof of Proposition 2.8.2.

A strategy σ is modified IS-time consistent, then there exists a weakly consistent belief $\mu(\sigma)$ such that for all sets of information sets $\mathscr{X}_1$ which are reached with positive probability under σ, we have

$$E^{IS}u(\sigma|\mathscr{X}_1) \geq E^{IS}u([\sigma;(\sigma',\mathscr{X}_1)]|\mathscr{X}_1),$$

for any behavioural strategy σ' and the weakly consistent belief $\mu([\sigma;(\sigma',\mathscr{X}_1)])$.

Assume $\mathscr{X}_1 = \mathscr{X}$, then

$$E^{IS}u(\sigma|\mathscr{X}_1) = E^{IS}u(\sigma|\mathscr{X}) = Eu(\sigma|\sigma).$$

Thus, the strategy σ is ex-ante optimal.

If a strategy σ is ex-ante time consistent, for any behavioural strategy σ' we have

$$Eu(\sigma|\emptyset) \geq Eu(\sigma'|\emptyset).$$

For any set of information sets $\mathscr{X}_1$, and any behavioural strategy σ'', we have

$$Eu(\sigma|\emptyset) \geq Eu([\sigma;(\sigma'',\mathscr{X}_1)]|\emptyset).$$

Denote $Z_0 = Z_{IS}(\mathscr{X}_1,\sigma) \cup Z_{IS}(\mathscr{X}_1,[\sigma;(\sigma'',\mathscr{X}_1)])$,

$$\sum_{z\in Z\setminus Z_0} p(z|\emptyset,\sigma)u(z) = \sum_{z\in Z\setminus Z_0} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)])u(z),$$

thus

$$\sum_{z\in Z_0} p(z|\emptyset,\sigma)u(z) = \sum_{z\in Z_0} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)])u(z).$$

For any behavioural strategy σ'', and a very small ε, construct $\sigma''_n = \frac{\varepsilon}{n}\sigma + (1-\frac{\varepsilon}{n})\sigma''$, then $\forall n \in \mathscr{N}$,

$$\sum_{z\in Z_0^n} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z_0^n} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)])u(z),$$

$$\sum_{z\in Z_{IS}(\mathscr{X}_1,\sigma)} p(z|\emptyset,\sigma)u(z) \geq \sum_{z\in Z_0^n} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)])u(z),$$

where $Z_0^n = Z_{IS}(\mathscr{X}_1,\sigma)\cup Z_{IS}(\mathscr{X}_1,[\sigma;(\sigma_n'',\mathscr{X}_1)]) = Z_{IS}(\mathscr{X}_1,\sigma)\cup Z_{IS}(\mathscr{X}_1,[\sigma;(\sigma'',\mathscr{X}_1)])$. It is easy to see that, when $n\to\infty$, $\sigma_n''\to\sigma''$. Then,

$$\sum_{z\in Z_{IS}(\mathscr{X}_1,\sigma)} p(z|\emptyset,\sigma) \geq \sum_{z\in Z_{IS}(\mathscr{X}_1,[\sigma;(\sigma'',\mathscr{X}_1)])} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)]).$$

If there exists a weakly consistent belief $\mu(\sigma)$ with σ and a weakly consistent belief $\mu([\sigma;(\sigma'',\mathscr{X}_1)])$ with $[\sigma;(\sigma'',\mathscr{X}_1)]$, we have

$$\frac{1}{p(\mathscr{X}_1|\sigma)} = \frac{1}{p(\mathscr{X}_1|[\sigma;(\sigma'',\mathscr{X}_1)])},$$

thus,

$$\frac{1}{p(\mathscr{X}_1|\sigma)}\sum_{z\in Z_{IS}(\mathscr{X}_1)} p(z|\emptyset,\sigma) \geq \frac{1}{p(\mathscr{X}_1|[\sigma;(\sigma'',\mathscr{X}_1)])}\sum_{z\in Z_{IS}(\mathscr{X}_1,[\sigma;(\sigma'',\mathscr{X}_1)])} p(z|\emptyset,[\sigma;(\sigma'',\mathscr{X}_1)]).$$

In other words,

$$E^{IS}u(\sigma|\mathscr{X}_1) \geq E^{IS}u([\sigma;(\sigma'',\mathscr{X}_1)]|\mathscr{X}_1), \text{ for any behavioural strategy } \sigma''.$$

The strategy σ is modified IS-time consistent.

Proof of Proposition 2.8.3.

For any optimal strategy $\sigma^\star$, construct a sequence of behavioural strategies such that, at the information set X where $p(X|\emptyset,\sigma^\star)>0$,

$$\sigma_n := \begin{cases} (1-\sum_{\sigma_X^\star(a)=0}\varepsilon(a))\sigma_X^\star(a) & \text{if } \sigma_X^\star(a)>0 \\ \varepsilon(a)\in(0,\frac{\varepsilon_0}{n}] & \text{if } \sigma_X^\star(a)=0 \end{cases}$$

where ε_0 is a very small number. At information set X where $p(X|\emptyset,\sigma^\star)=0$, the behavioural rule at X can be arbitrary.

Model the decision problem Γ as a multi-agent game $(A(\Gamma), \frac{\varepsilon_0}{n})$ where the decision maker at the set of information sets $\mathscr{X}_1$ and all the information set $X \prec' \mathscr{X}_1$ is an agent with its utility function $E^{IS}u(\mathscr{X}_1)$and he at other information set X' as each individual agent with its utility function $E^{IS}u(X')$. Then, by Kakutani's fixed point theorem, there must be a Nash equilibrium, $\sigma_n^\star$ in each $(A(\Gamma), \frac{\varepsilon_0}{n})$. Since Γ is compact, there must be a subsequence of $\sigma_{n_k}^\star$ such that $\lim_{n_k \to \infty} (\sigma_{n_k}^\star, \mu_{n_k}^\star) = (\sigma, \mu)$ where $\mu_{n_k}^\star$ is the weakly consistent belief of $\sigma_{n_k}^\star$. It is easily to see σ is strongly consistent with μ and σ and σ is outcome equivalent to $\sigma^\star$.

Chapter 3

Sequential Equilibrium in Games with Imperfect Recall

3.1 Introduction

In the analysis of time consistency in decision problems, sequential rationality can be a criterion to evaluate a strategy profile in games. A strategy profile is **sequentially rational** if, given a system of beliefs and the strategies of other players, at any information set, the expected payoffs assigned to each player's strategy are higher than any other strategies. In other words, the strategy of each player from a sequentially rational strategy profile is conditionally optimal over all available strategies at any information set. Correspondingly, sequential equilibrium might be the most common solution concept that is discussed in games. However, the discussions are restricted to games with perfect recall. In this paper, we only discuss finite games with complete information.

A player is said to have **perfect recall** if, at any instance where he moves, he remembers all his past actions and knowledge has been acquired. Otherwise, the player is said to have imperfect recall. A game is said to have perfect recall if every player has perfect recall. There is a special case of imperfect recall, in which a player might forget whether he has made some moves at one of the instances where he moves consecutively.

Those consecutive histories where the player moves and does not know whether he has passed some of them belong to one information set of him. The available behavioural rules at each history in that information set are the same which is one of the conditions that make a player do not know which node in the information set he is currently located. A path of the game tree which passes this information set might contain more than one history from that information set. In that case, we call the information set presents **absentmindedness**. The name comes from the example of absentminded driver proposed by Piccione & Rubinstein (1997a). We use PR to represent the paper henceforth. Imperfect recall is a special case of imperfect information. Games with perfect information require that each player is informed of all the events that occurred previously at any instance where he moves.

Besides absentmindedness, there are many types of imperfect recall. For example, like in the example of absentminded driver paradox (see Figure 3.1), at the first intersection, the driver knows he is there, but he would assign the probability to being at the first intersection if he moves to the second intersection in the future. Such a structure of imperfect recall is very hard to model. However, in this paper, we discuss three main types of imperfect recall. The first type (see Figure 3.4) is that a player forgets his previous actions. The second type (see Figure 3.10) is that a player forgets his previous acquired. As shown in Figure 3.10, another player A moves at the red node. The player knows the action of player A at e_1 or e_2 but forgets it when he moves to the information set X. Player A could also be the chance player. The third type (see Figure 3.11) is that a player presents absentmindedness. As shown in Figure 3.11, a player (the driver) moves at information set X_D and the path from e_0 to z_1 cuts the information set X_D twice, i.e., nodes e_0 and e_1 belong to that path.

There are many differences between games with perfect recall and imperfect recall. In games with perfect recall, the beliefs at a non-singleton information set are determined by the previous actions of players. However, with the presence of imperfect recall, a player's belief at an information set might be also determined by his current strategy. For example, in the absentminded driver paradox, the decision maker's belief at the only information set is determined by his behaviour rule at that information set. As it has been analysed in the previous chapter, the decision maker's choice over the available strategy set determines the strategy and the belief simultaneously.

Besides, in games with perfect recall, a behavioural strategy of a player has an outcome-equivalent mixed strategy, and vice versa. There must exist a Nash equilibrium in mixed strategy in a finite game of perfect recall. Associated with the fact of outcome equivalence between mixed and behavioural strategies, a Nash equilibrium in behavioural strategies must exist in a finite game with perfect recall. However, allowing imperfect recall, Nash equilibria may not exist in a finite game. Wichardt (2008) proposes an example in which no Nash equilibrium in behavioural strategies. Lambert et al. (2019) extend the absentminded driver paradox in PR to a two-player game, named absentminded driver and the policeman (see Figure 3.11), to show there is no Nash equilibrium in behavioural strategies in that example. They also propose another version of extensive game of absentminded driver and explain there exists no sequential equilibrium and perfect equilibrium, although a Nash equilibrium with noncredible threat exists in that game. The nonexistence of Nash equilibrium, sequential equilibrium or perfect equilibrium in a finite game with imperfect recall motivates them to explore alternative equilibrium notions instead of the classic one-self approach. Thus, they term those classic solution concepts in a multiself approach. Hillas et al. (2020b) state that there might not be any Nash equilibrium in mixed strategies in a nonlinear game. Actually, we have not found an example to verify it. However, it reminds us to discuss solution concepts in a more general class of strategies in a finite game with imperfect recall.

In this paper, we apply the sophisticated recursive calculation rule to a finite game with imperfect recall. For simplicity, the game is assumed to have complete information and the other players have perfect information except for one player who has imperfect recall. Even for the player who has imperfect recall, we assume he has perfect information about the other players' strategies. In fact, it does not matter whether there is more than one player who presents imperfect recall as long as the players have perfect information about the other players' strategies. The imperfect recall is an internal imperfection. The third party cannot observe it directly but might deduce it through the player's "irrational" behaviours. If we apply the sophisticated recursive calculation rule to the finite game, the conditional expected payoff function for the player with imperfect recall at an information set becomes IS expected payoff function.

A sophisticated player knows he will present imperfect recall in the future execution of the game. At that information set with imperfect recall, the player does not know which node he is currently located. The player has to choose a strategy based on the belief system he forms which is consistent with the strategies of the other players and the strategy he executes until the current node. Then, he might consider some outcomes which will never be achieved by any strategy due to his imperfect recall. Although the player at the current node might be able to know which outcomes are impossible to realise at that information set with imperfect recall, he would rather behave as if he does not know the information about the outcomes. Then, it is possible for him to choose a strategy at the current node which is also conditionally optimal when he reconsiders at that information set since he currently considers the game in the same way as if he moved to the future information set. As what has been discussed above, we would like to examine the sequential rationality of a strategy. In our paper, we use IS expected payoff function as the standard to examine the sequential rationality of a strategy.

The main motivation is to see whether the sophisticated recursive calculation rule can be applied to a finite game. Furthermore, if the resulted IS expected payoff function is consistent with that in a finite decision problem, it is interesting to see whether a strategy is ex-ante optimal if it is sequentially optimal, given the strategy profile of the other players.

The IS expected payoff function at the information set which contains the first node where a player acts coincides with the ex-ante expected payoff function in conventional game theory. Thus, an IS sequential equilibrium must be a Nash equilibrium. The existence of an IS sequential equilibrium indicates that there exists a strategy profile such that, conditional on the perfect knowledge about the other players, no player has a conditionally more beneficial strategy than the strategy described in the equilibrium strategy profile when any player reconsiders at any information set, if the player with imperfect recall is sophisticated.

Besides the IS sequential equilibrium, we also discuss the inequivalence between behavioural strategies and mixed strategies in non-linear finite games, such as absentminded driver paradox. Therefore, we have to introduce the concept of general strategies. A general strategy involves in two-stage randomisation over pure strategies: ex-ante and during execution. Therefore, it includes both behavioural strategies and mixed strategies. The analysis of imperfect recall in a finite game should be discussed over general strategies.

The rest of this paper is organised as follows. The explanation of a player and a strategy is discussed in the section 3.3 and 3.4. The sophisticated recursive calculation rule and IS expected payoff function are described in section 3.6. In section 3.7, we define IS sequential equilibrium which satisfies sequential rationality in terms of IS expected payoff function. The chain of inclusion among solution concepts in games with imperfect recall is the same as is in games with perfect recall.

3.2 Relevant Literatures

The games discussed in this paper allow both imperfect recall and the game cuts a player's information sets more than once. Kuhn (1950, 1953, 2009) proposes that, in a game with perfect recall, every player remembers everything he knew and how he behaved in the past at each occasion he moves. A game which does not satisfy the above condition presents imperfect recall. However, in his definition of extensive games, each play of a game cuts each information set at most once. In other words, there is at most one node of every information set along a path in a game. Isbell (1957, 1959) extends the scope of games by allowing more than one occurrence of each information set in any play of a game. He calls the games in which each information set is passed at most once linear games, otherwise where the requirement is not met nonlinear games.

The absentminded driver paradox makes the interpretation of games with imperfect recall in the spotlight and many other works (Aumann et al. 1997a,b, Battigalli 1997, Gilboa 1997, Grove & Halpern 1997, Halpern 1997, Lipman 1997, Piccione & Rubinstein 1997b) follows PR. The explanation of paradoxical facts in the example of absentminded driver accounts for a great proposition in their analysis. In this paper, we are going to apply the framework of the sophisticated recursive calculation rule to general extensive games and discuss the relevant solution concepts.

In games with perfect recall, a Nash equilibrium must exist if the strategy set for each player is nonempty, convex and compact and the utility function of each player is continuous over the set of strategy profile sets and quasi-concave in his strategy set. However, according to Wichardt (2008), there might exist no Nash equilibria in behavioural strategy profile if the game presents imperfect recall. Furthermore, with perfect recall, there exists a chain of inclusion that a sequential equilibrium must be a Nash equilibrium and a perfect equilibrium must be a sequential equilibrium. Kline (2005) demonstrates the inclusion chain breaks when the assumption of perfect recall is not valid. Therefore, it is easy to see that the existence of different solution concepts and the relation among them might be totally different when an extensive game has perfect recall from when it does not.

The fact that some crucial properties in games of perfect recall are not valid anymore if the assumption is dropped might cast doubt on the correctness of classic solution concepts on the occasion of imperfect recall. Halpern & Pass (2016) propose their definitions of quasi-perfect equilibrium, ex-ante sequential equilibrium and interim sequential equilibrium based on behavioural strategy mixtures. They explain that the partial order among information sets in perfect recall might not exist in games with imperfect recall. The equilibrium concepts require a player's strategy should be the best response at any set of information sets among strategies which do not affect the probability to the terminal nodes which cannot be reached with positive probability starting from that set of information sets. Hillas & Kvasov (2020a) also present their definition of equilibrium concepts which follow the spirit of backward induction and are comparative in linear games with imperfect recall. Not only do they consider the deviations at multiple information sets but also will each player's belief at an information set depend on the player's current strategy and each player's belief over other players' strategies probably differ from his ex-ante expectation. They also define the general strategy which assigns a probability distribution over behavioural strategies.

Compared to Halpern & Pass (2016) as well as Hillas & Kvasov (2020a), which assume every player is able to coordinate his behaviours at various instances where he moves, Lambert et al. (2019) develop the multiself approach with imperfect recall. He adopts the formation from Gilboa (1997) and extends it to extensive games. In their paper, a phantom strategy specifies a probability of behaviours at a particular node in a non-singleton information set instead of the identical behavioural rule at different nodes in the same information set as required in the classic framework. It models a player as multiple selves with the identical preference who make decisions independently. At an information set, the histories which have the same depth are assumed to be a separate self and one-shot deviation is mainly discussed. The authors define every equilibrium concept in the notion of multiselves and the corresponding inclusion chain is consistent with that in games with perfect recall.

In this paper, we also follow Halpern & Pass (2016), Hillas & Kvasov (2020a) and discuss games with imperfect recall in one-self approach in which players have some ability of coordination across information sets but in the framework of IS expected payoff function.

3.3 What Is a Player?

A player seems to be an entity who can formulate an optimal plan for a game before it starts and implement the chosen plan during the execution. He could well coordinate across the choices at different instances to achieve his highest payoffs. It reflects his control power over actions. However, solution concepts, such as sequential equilibrium, provide the player with a channel to evaluate his ex-ante chosen strategy during the execution of the game. The control power of a player and his evaluation criterion at the instance where the player reevaluates is important. The evaluation criterion is the conditional expected payoff function. A strategy is conditionally optimal if the player cannot increase his expected payoffs by deviating to a different strategy. In our paper, the classic expected payoff function is replaced by IS expected payoff function as the criterion when the player does the reconsideration. As for the control power of a player, researchers hold two contradicted viewpoints. Some of them, such as Piccione & Rubinstein (1997a), Gilboa (1997) and Lambert et al. (2019), argue that the player's control power is restricted to his immediate behaviour when he reconsiders. The main reason behind it is that the player who has imperfect recall might not remember the updated strategy if he deviates at an information set. For example, see Figure 3.10, if the player at e_1 changes his strategy, he might be unconscious of the previous deviation when he is at X since he even forgets which node he has passed to reach X. It is practical to assume that a player notices the potential deviation from the initial chosen strategy only if he reaches an information set which should not be reached by following his initial strategy. In that case, the player is modelled to be multiple agents who work as a team with identical preferences. Each agent moves once in a game and no communication among them. In a linear game, it is equivalent to analysing the game in the agent form. On the contrary, in a nonlinear game, it might involve different actions at different nodes in an information set. However, the other researchers, such as Segal (2000), Dimitri (1999), Dimitri (2009) and Battigalli (1997), believe, in spite of the lack of perfect recall, a player could coordinate his choices and fully formulate the strategy plan for the rest of play. In that case, one would model a game in the one-self approach.

In fact, the above two assumptions of players are extremely opposite regarding the ability of coordination and control power. The player (in one-self approach) has full control over the moves at the histories of which the current history is a subhistory and could coordinate across multiple histories. In the meanwhile, the player (in multiselves approach) could be regarded as multiple agents whose control power of behaviours are restricted to their immediate moves. They cannot communicate with each other, and naturally the player has no ability of coordination. However, in our framework, we assume that a player has control over the moves at later information sets, but at the current information set where the player is, he only has control over his immediate move. The reason why we assume limited control power at the current information set is that the player is defined as lose of memory about his previous actions or previously acquired information. It causes the decision maker cannot know which node in the current information set he is now located. If the current information set presents no absentmindedness but the other types of imperfect recall, such as example 1 (see Figure 3.4) and example 2 (see Figure 3.10). The planner moves out the current information set by one-shot action. If the current information set presents absentmindedness, the nature of it makes it impossible to record completely which nodes he has passed and the strategies implemented at those past nodes. It is different from the situation that two histories belong to different information sets. For example, if h_1 and h_2 belong to two information sets and h_1 is a subhistory of h_2, assume that a player changes his strategy at h_1 and when he moves to h_2, he is able to deduce the updated strategy which is chosen at h_1 by the assumption of potential reconsideration at the information set containing h_1. Notice that, a behavioural rule is different from an actual action if the behavioural rule is not deterministic. For example, as shown in Figure 3.4, being at the information set K, a player might know his behavioural rule at e_1, i.e., the probability measure of actions L and R, but forgets which action he takes at e_1. Therefore, it is possible that players know the updated strategy whenever they move to a new information set with the help of an external device or the third party who is the observer of the game but forget the previous actions. Thus, the alteration of strategy occurs either at histories where the player does not even know whether he has passed or at histories in other paths of the game. In both cases, not only the updated strategy but also the signal that he has reevaluates the game cannot be conveyed to other histories in the same information set.

It is easy to see that the assumption of a player's control power in our framework is in between the two extreme assumptions. Although a player is assumed to have limited control power of his immediate behaviour at a non-singleton information set, the player knows, that once he reconsiders at an information set, he will be reconsiders at every node in that information set. He expects that he would solve the same optimality question at other histories in that information set, and thus, the same optimal behavioural rule should be extracted when he maximises the expected payoff function. In other words, a player expects to also behave the same at other histories in the information set.

There is a major different assumption between our framework and the models developed by other works. Most works on games with imperfect recall assume either each player reconsiders once (at a particular history in an information set) or reconsiders at an information set (each history in the information set). Therefore, a player does not need to think about whether his switched strategy satisfies conditional optimality at an information set that he will pass in the future if he deviates from the initial strategy. However, in our framework, a player will be stopped at any time and ask him to reconsider at an information set. It is possible that he is stopped only once or at each information sets. The players do not know where he will be asked to reconsider. It is the reason why, in our framework, a player takes how he would think if he also reconsiders at future information sets into consideration when he reevaluates at the current one. The framework illustrates the inspirit of backward induction. A player knows how he will behave at the very end of the game based on any given strategy. Then, the player at the second to the last step of the game incorporates how he evaluates at the end of the game into his current evaluation of the game. Even though we do not know how a player with imperfect recall remembers the updated strategy, we assume a player is capable of coordinating and implementing a coherent strategy.

3.4 What Is a Strategy?

Morgenstern & Von Neumann (1953) develop the concept of strategies. They describe a strategy as a complete contingent plan, or behavioural rule, which specifies how the player should move at each possible circumstance in which he is called upon to move. A strategy should specify a player's action at some node which might not be reached during the actual execution of the game. It should work as an instruction book which is written down prior to the play. Thus, a third party who acts as the action representative of a player could play the game without the player by following the instruction book. Then, Kuhn (1950) defines a pure strategy as a function which maps a choice of action to an information set, even those information sets that are precluded by the strategy itself. He thinks it is a deficiency of the definition since it is too redundant. Then, he simplifies the definition by defining an equivalent class of strategy. Two pure strategies are equivalent if, for any strategies played by other players, the two strategies induce the same probability measure over different outcomes. He manifests that we actually mean an equivalent class of strategies when we talk about a particular pure strategy. In fact, when we evaluate the weak notion of time consistency of a strategy, if a strategy is time consistent then its equivalent strategy is also time consistent. However, Rubinstein (1991) rejects the idea of equivalent class. He indicates that a strategy should include not only his action plan but also how the other players' beliefs in the event that he does not follow the plan. In this paper, we do not discuss too much about the equivalent class of a strategy since the solution concepts of sequential equilibrium and perfect equilibrium require strategies that can reach any information set in the game.

A mixed strategy is randomisation over pure strategies. The randomisation occurs prior to the play of the game. Once the game starts, a player who implements a mixed strategy actually implements a deterministic action at each node where he moves. Thus, when a player considers a deviation from the initial strategy at some instance in the environment of mixed strategies, he could only deviate to any pure strategy. In fact, in the definition of one-shot deviation principle, a strategy is conditional optimal if no other pure action reaches higher expected payoffs than the action pointed out by the strategy. Another main character of pure strategies and mixed strategies is that it shows a player's perfect ability of coordination across information sets. In other words, we should analyse a game in terms of mixed strategies if we explain a player as an entity with the ability of coordination. On the contrary, a behavioural strategy is a function which assigns a probability measure over available actions to each information set. The probability measures at different information sets are independent of each other. Thus, for a behavioural strategy, it does not matter whether a player does the randomisation ex-ante or during execution. A behavioural strategy allows freedom of randomisation at each information set while presenting little ability of coordination of actions at multiple information sets. Thus, solution concepts, such as sequential equilibrium, are defined in terms of behavioural strategies since they present notion of interim. However, in a nonlinear game, at an information set presenting absentmindedness, a behavioural strategy works as a pure strategy at an information set without absentmindedness. At an information set without absentmindedness, there must exist a deterministic action which is weakly dominated any combination of available actions. However, at an information set presenting absentmindedness, there could be a combination of actions which attains higher expected payoffs than any deterministic action. For example, in the absentminded driver paradox, assume that the driver receives the payoff of 0 if he exits the highway at the first intersection, payoff of 4 if he exits at the second intersection and payoff of 1 if he keeps on driving on the highway to the end. The optimal strategy at the information set is to continue on the highway with the probability of $2/3$ (expected payoff of

8/3) and it attains higher expected payoffs of the deterministic action of “CONTINUE" at both intersections (payoff of 1) and the deterministic action of “EXIT" at both intersections (payoff of 0). If a player reconsiders at an information set with absentmindedness, he might deviate to a new probability measure over available actions instead of a deterministic behaviour. Thus, the definition of solution concepts in nonlinear games should require a strategy are conditionally optimal over any combination of available actions instead of any deterministic available action.

Therefore, to analyse games with imperfect recall, we would like to develop distributions over behavioural strategies. By following Mertens et al. (2015) and Hillas et al. (2020b), we call such strategies general strategies. In other previous works, such as Selten (2020), general strategies are also called behaviour strategy mixtures. A **general strategy** involves two types of randomisation: ex-ante and during execution. It is easy to see that behavioural strategies and mixed strategies are special cases of general strategies. By allowing only ex-ante randomisation of a general strategy, we could get a mixed strategy. By allowing randomisation only in the course of a game, we could get a behavioural strategy. Naturally, pure strategies are included in general strategies. We could interpret a player who implements a general strategy that he randomises over behavioural strategies at the beginning of the game and choose one. During execution, the player follows the chosen behavioural strategy throughout the game. Although the player has the capability to do randomisation at each information set, he should commit to the behavioural rule described in the ex-ante chosen behavioural strategy (Halpern & Pass 2016). The notion of equivalent strategies could be extended in terms of general strategies and nonlinear games. Two strategies of a player are defined to be Kuhn equivalent if, given any general strategy profile of the other players, the two strategies induce the same measures over the outcomes. According to Kuhn's theorem, for any mixed strategy, there is a Kuhn equivalent behavioural strategy, It is proposed and proved by Kuhn (1950, 1953) if the player has perfect recall. For any behavioural strategy, there is a Kuhn equivalent mixed strategy, if the game is linear. It is formally developed by Isbell (1957, 1959).

3.5 Notations and Definitions

In a finite extensive n-player game Γ, define $N = \{0, ..., n\}$ the set of players, where player 0 represents the chance player and A the set of available actions. Denote a^i an action of player i and A^i the set of available actions of player i. There are acyclic connected nodes with a unique initial node as the root. The directed connection from the root node to a node which is only connected to another node, we call it a terminal node, is defined as a directed branch. A terminal node represents an outcome of the game. Every node can also be alternatively expressed as a sequence of all the previous actions that leads the game from the root node to the current node. We call the sequence a history $h \in H$. H is the set of histories, and an empty sequence $\emptyset$ which present the root node also be included. Denote $h \in h'$ if $\exists a_1, ..., a_k$ such that $h' = (h, a_1, ..., a_k)$. We call h **precedes** h' or h is a subhistory of h' and write $h \prec h'$. Denote $A(h)$ the available actions at the history h. Denote Z the set of terminal nodes, thus $Z := \{h \in H | \forall a \in A, (h, a) \notin H\}$. Define the **immediate predecessor function** $P(\cdot) : H \setminus Z \to H$ such that if $P(h') = h$, we have $h' = (h, a), a \in A$. We have $P_{k+1}(h') = P(P_k)(h') = h$ if $h' = (h, a_1, ..., a_{k+1}), a_1, ..., a_{k+1} \in A$. We say there is a **path** from history h_1 to h_2 if $\exists l$ such that $h_1 = P_l(h_2)$. We use $Z(h)$ to express the set of terminal histories to which there is a path from h. In this paper, we do not particularly distinct between a history and a node. Define $\alpha : H \setminus Z \to N$ the **player function**, it assigns a distinct player $i, i \in N$ to every node expect terminal nodes. Denote $\rho : H \setminus Z \to \Re([0, 1])$ the players' beliefs on the actions of the chance player. Denote the utility function of player i $u_i : Z \to \Re$, it assigns the payoff of the player i to each terminal node. It represents players' preferences over outcomes.

The history set H can be partitions into several subsets. Denote X^i an information set where the player i operates, if $\forall h, h' \in X^i$, then $\alpha(h) = \alpha(h') = i$, $A(h) = A(h') = A(X^i$, where $A(X)$ the available actions at information set X. We use $Z(X)$ to express the set of terminal histories to which there is a path from a history in X. In perfect recall, we have $Z(X) \supseteq Z(X')$ if $X \prec X'$. However, it is not the case in games with imperfect recall in the relation $\prec'$. For example, see Figure 3.10, $Z(\{e_2\}) = \{z_2, z_5, z_6\}$, $Z(X) = \{z_3, z_4, z_5, z_6\}$ and $\{e_2\} \prec' X$. Also denote X_h the information set containing the history h. When a player reaches an information set, assume that he does not know which node in the information set he is currently located. Denote $\mathscr{X}$ the collection of all information sets and $\mathscr{X}_i$ the collection of all information sets where the player i moves. Define the partial order $\prec$ over information sets. Write two information sets $X, X' \in \mathscr{X}$, $X \prec X'$ if $\forall h' \in X', \exists h \in X$ such that $h \in h'$. Now extend the immediate predecessor function to information sets. Define $X = P_{IS}(X')$, if $\forall h' \in X', \exists h \in X, a \in A(X)$ such that $h' = (h, a)$. Also define the relation $\prec'$ such that, if for two information sets $X_1, X_2 \in \mathscr{X}$, we have $X_1 \prec' X_2$, then $\exists h_2 \in X_2, h_1 \in X_1$ such that $h_1 \prec h_2$. In the previous section, we use the phrase “a later information set" if the information set is reached later than the current information set. Being at an information set X, we are talking about those information sets X' satisfying $X' \prec' X$ or $X' \prec X$, depending on which relation is established in the game being analysed, if we use the phrase similar to “at later information sets", “ at later stage of the game", “after X", and so on.

With imperfect recall, there could be two information sets $X, X' \in \mathscr{X}$ such that $X \prec' X'$ and $X' \prec' X$. For example (see Figure 3.12), assume that player 1 moves at information set X_1 and player 2 moves at information set X_2. It is actually the Gilboa formation of absentminded driver paradox which formats a decision problem with absentmindedness into a two-agent game form. We could frame the example in the viewpoint of player 1, for example. As the player 1, I will not be informed about whether player 2 has moved or not. Thus, node e_1 and e_4 are identical to me. When it is my turn to move, I do not know which node I am currently at. We call two information sets X_1, X_2 are **cross-branch information sets** if $X_1 \prec' X_2$ and $X_2 \prec' X_1$. We call information sets $X_1, ..., X_K, K \geqslant 2$ are cross-branch information sets, if $\forall i, j \in \{1, ..., K\}$, we have $X_i \prec' X_j, X_j \prec' X_i$. Now define the **immediate predecessor correspondence** P'_{IS} on the relation $\prec'$. Notice that, for any information set, there might be more than one predecessor information set. For example, see Figure 3.10, the information sets $\{e_1\}$ and $\{e_2\}$ are both predecessor information sets of information set K. Thus, P'_{IS} is a correspondence instead of function. Denote $\mathscr{Y}$ the union set of $\mathscr{X}$ and $\{z\}, z \in Z$. Then, $Y \in \mathscr{Y}$ is either an information set $X^i, i \in N$ or a singleton set containing a terminal node $z \in Z$. We say there is an **IS-path** from information set X to Y if $\exists X_j, j = \{1, ..., l\}$ such that $X \in P'_{IS}(X_1)$, $X_1 \in P'_{IS}(X_2)$,..., $X_l \in P'_{IS}(Y)$. Notice that $P'_{IS}(\{z\}), z \in Z$ is a function rather than a correspondence. For any information set $X \in \mathscr{X}$ and $Y \in \mathscr{Y}$, there might be more than one IS-path. For example, see Figure 3.5, from the information set $\{e_0\}$ to information set X, an IS-path is $\{e_0\}, \{c\}, \{e_1\}, X$ and the other one is $\{e_0\}, \{c\}, \{e_2\}, X$. Additionally, also define $P'^i_{IS} : \mathscr{X}^i \to 2^{\mathscr{X}^i}$ the correspondence which indicates the immediate predecessor information set in $\mathscr{X}^i$ for player i. Correspondingly, also define $\mathscr{Y}^i := \mathscr{X}^i \cup \{z, z \in Z\}$. We say there is an ***i*-IS-path** from information set X^i to Y^i if $\exists X^i_j, j = \{1, ..., l\}$ such that $X^i \in P'^i_{IS}(X^i_1)$, $X^i_1 \in P'^i_{IS}(X^i_2)$,..., $X^i_l \in P'^i_{IS}(Y^i)$. Thus, in a game with imperfect information, there must be only one i-IS-path from X^i to Y^i but several IS-paths from X^i to Y^i. Denote $Z_{IS}(Y)$ the set of terminal nodes such that $\forall z \in Z_{IS}(Y)$, there is an IS-path from Y to $\{z\}$. For simplicity of notation, if $Y = \{z\}$, we have $Z(Y) = \{z\}$. Then, $Z_{IS}(X) \supseteq Z_{IS}(X')$ if $X \prec' X'$. For those terminal nodes $z \in Z_{IS}(X')$, there is an IS-path from X' to $\{z\}$. There is an IS-path from X to X' when $X \prec' X'$. Thus, there is an IS-path from X to $\{z\}$, $z \in Z_{IS}(X)$.

Now define the system of strategies. Denote $s^i \in S^i$ the pure strategy of player i and S the set of pure strategy profiles. Denote $\pi^i \in \Pi^i$ the mixed strategy of player i and Π the set of mixed strategy profiles. Denote $\sigma^i \in \Sigma^i$ the behavioural strategy of player i and Σ the set of behavioural strategy profiles. Denote $\omega^i \in \Omega^i$ the general strategy of player i and Ω the set of general strategy profiles. Denote $K^{-i} := \times_{j \neq i} K^j$, where $K \in \{S, \Pi, \Sigma, \Omega\}$. The ex-ante expected payoff function for player i with the general strategy profile ω is

$$Eu_i(\omega|\emptyset) = \sum_{z \in Z} p(z|\emptyset, \omega) u_i(z).$$

A player i is said to have perfect recall, if for any two histories e_1, e_2 in any information set of player i and there is a node e'_1 at which player i acts, such that $e'_1 \prec e_1$, there is also a node e'_2 in the same information set as e'_1 such that $e'_2 \prec e_2$, and the action of player i at e'_1 on the path to e_1 is the same as the action of player i at e'_2 on the path to e_2 (Kuhn 1950). A game is said to have perfect recall if all players have perfect recall. Define $z(s) \in Z$ the terminal node reached by the pure strategy profile s. Kuhn (1953) defines a game in terms of sets of terminal nodes. We interpret it by the language consistent with this paper. A player i has perfect recall, for all pure strategy of player i and all information set X^i, if there exists a history $h \in X^i$ such that $\exists s^{-i} \in S^{-i}, h \prec z(s^i, s^{-i})$, then $\forall h' \in X^i, \exists s'^{-i}$ such that $h' \prec z(s^i, s'^{-i})$. A game has perfect recall if all players have perfect recall. Later, in 2003, Kuhn demonstrates the two definitions are equivalent in his notes (Kuhn 2009). Both definitions describe that any two histories in an information set of a player with perfect recall are reached by the same previous actions of that player and different previous actions by other players. However, absentmindedness is allowed in our discussion, a game which satisfies those conditions might contains information sets presenting absentmindedness. An extra condition should be added to well define perfect recall in a general finite game with complete information.

Definition 3.5.1 *A player i is said to have **perfect recall**, if for any information set X^i and any two histories $e_1, e_2 \in X^i$,*

1) we do not have $e_1 \prec e_2$ or $e_2 \prec e_1$;

2) if whenever there is a node e'_1 in an information set where player i operates and $e'_1 \prec e_1$, there is also a node e'_2 in the same information set as e'_1 such that $e'_2 \prec e_2$ and the action of player i at e'_1 on the path to e_1 is the same as the action of player i at e'_2 on the path to e_2.

Definition 3.5.2 *An information set X is said to present* ***absentmindedness****, if $\exists e_1, e_2 \in X$, such that $e_1 \prec e_2$. A player is said to have absentmindedness if one of his information set presenting absentmindedness.*

Now we define the notation system relevant to probability distribution. Define $p(h'|h,\omega)$ the probability of reaching the history h' from the history h by playing a general strategy ω. If $h = \emptyset$, it presents the probability of reaching the history h' from the beginning of play. Denote $\hat{X}$ the **upper frontier** of the information set X and $\hat{X} := \{h \in X |$ there does not exist any $h' \in X,$ such that $h' \in h\}$. Define the probability of reaching information set X is

$$p(X|\emptyset,\omega) = \sum_{h\in\hat{X}} p(h|\emptyset,\omega).$$

For a game Γ, define a **belief system** $\mu : \mathscr{X} \to [0,1]$ is a function which assigns a probability measure $\mu(X)$ over histories in X for any information set X in Γ. For a history $h \in X$, $\mu(h|X)$ is the probability of currently being at history h conditional in the information set X. Then, we require $\sum_{h\in X} \mu(h|X) = 1$. A belief system μ is defined to be **weakly consistent** with a general strategy ω, if for any information set X which is reached with positive probability and every history $h \in X$, we have

$$\mu(h|X,\omega) = p(h|\emptyset,\omega) / \sum_{h'\in X} p(h'|\emptyset,\omega).$$

A behavioural strategy σ^i is called to be a **completely mixed** behavioural strategy, if for any information set X^i and action $a \in A(X^i)$, σ^i assigns positive probability to implementing the action a. A general strategy ω^i is completely mixed if every behavioural strategy in its support is completely mixed. A belief system μ is **consistent** with a general strategy profile ω, if there is a sequence of completely mixed general strategy profile $\omega_1, \omega_2, ...$ and their weakly consistent belief systems $\mu(\omega_1), \mu(\omega_2), ...$ converging to the general strategy ω and its weakly consistent belief system $\mu(\omega)$. Alternatively, we could say an assessment (ω,μ) is consistent if the belief system μ is consistent with the general strategy ω.

Although in most works, a belief system is defined to assign a probability distribution over histories in all information sets in a game, Hillas & Kvasov (2020a) propose that, instead of the probability distribution, it is more efficient to define the belief of a player is about what strategies the other players are playing. Based on the strategies and the probabilities of playing those strategies of the other players, the probability distribution over histories at the current information set is generated and the player chooses his own strategy according his beliefs and the induced probability distribution. It avoids the redundancy to choose a strategy based on any strategy profile of the other player. However, taking the potential mistakes into consideration, we follow the way that most works do to define a belief system as the probability measures over histories at all information sets.

3.6 IS Expected Payoff Function

In a game, players are assumed to do their best response to the strategy profile of the other players. Before the game starts, based on the strategy profile of the other players ω^{-i}, the player i should choose a general strategy ω^i which maximises the ex-ante expected payoff function of player i,

$$Eu_i(\omega^i|\emptyset, \omega^{-i}) = \sum_{z \in Z} p(z|\emptyset, \omega) u_i(z).$$

By implementing a general strategy profile, every player chooses a general strategy at the beginning of a game. However, during execution, only one behavioural strategy from the support of the chosen general strategy is executed. Every player is assumed to know his latest strategy being played but might not remember the results of execution. In this paper, we model sophisticated players, in the sense that the player takes how he evaluates the game conditional at each information set into consideration when examining the sequential rationality of a behavioural strategy in middle of the game. At an information set, some terminal nodes which cannot be reached by any completely mixed general strategy are also considered if the player thinks he might reach those nodes at a future information set due to imperfect recall. When the player has perfect recall, it does not matter whether he is sophisticated since one-deviation property satisfies. It is enough for a player to consider his immediate behaviour. However, if a player presents imperfect recall, his previous actions, or even current action, determines his beliefs over an information set, and thus have effects on his decision on the behavioural rule of the following game. We will explain what the following game means later. Then, a sophisticated player is assumed to notice the fact and chooses a strategy which is optimal for himself across information sets.

3.6.1 IS Expected Payoff Function at an Information Set

In classic theory, a player at a history evaluates his conditional probability to each terminal node. In the presence of imperfect recall, at a non-singleton information set, a player is assumed to have control at his immediate action and the behavioural rule at different information sets which might be passed by playing a strategy from the current information set. Furthermore, the player knows not only the fact that he does not know which node in the information set he is currently located but also the evaluation of the expected payoffs at an information set is the same as any section of the information set. Thus, a sophisticated player understands his immediate action would lead him to an information set instead of a particular node, and the expected payoffs he could gain from his behaviour is the modified expected payoffs of the information set he reaches. Thus, we actually calculate the conditional probability from an information set to a terminal node through an IS-path. In games with perfect recall, it is equivalent between a path and an IS-path and, for any two information sets X_1, X_2 such that there does not exist $X_1 \prec X_2$ or $X_2 \prec X_1$, a terminal node should be included in at most one set of $Z(X_1)$ and $Z(X_2)$. However, with imperfect recall, for two information sets X_1, X_2 such that there does not exist $X_1 \prec' X_2$ and $X_2 \prec' X_1$, a terminal node might be included into both $Z_{IS}(X_1)$ and $Z_{IS}(X_2)$. For example, see Figure 3.10, we have $z_3 \in Z_{IS}(\{e_1\})$ and $z_3 \in Z_{IS}(\{e_2\})$. When considering at an information set, there is at most one IS-path which contains the path to a terminal node. Because a player does not know his precise current history at a future information set with imperfect recall, a sophisticated player notices the fact and keeps the his future ignorance about the node he is at that information set when he considers now. Then, the player behaves as if he does not know which IS-path contains the real path to any terminal node and without any prior. For the set containing a terminal node, each IS-path should be assigned with equal probability to be the one containing the path.

We develop the calculation of conditional expected payoffs into two steps. In the first step, the conditional probability of reaching each terminal node by an i-IS-path is evaluated. The conditional probability is also modified by the attitude of the player who acts at that information set to the expected payoffs at his next active information set. Define $\Phi^i(\cdot|X^i) : \Omega \to \times_{|Z|}(\Re \times Z)$ the **attitude modified IS-path probability function**, where $|Z|$ is the number of terminal nodes. Thus,

$$\Phi^i((\sigma^i,\omega^{-i})|X^i) = f_{X^i,(\sigma^i,\omega^{-i})}(\sum_{a\in A(X^i)} \sigma^i_{X^i}(a) \sum_{h\in X^i} \mu(h|X^i,(\sigma^i,\omega^{-i})) \sum_{\substack{h'\in \hat{Y}^i \\ X^i\in P'^i_{IS}(Y^i)}} p(h'|(h,a),\omega^{-i})\cdot \tag{3.1}$$
$$\lambda(h,a|(\sigma^i,\omega^{-i}),X^i)\Phi^i((\sigma^i,\omega^{-i})|Y^i))$$

where

$$\lambda(h,a|(\sigma^i,\omega^{-i}),X^i) := \begin{cases} 1 & \text{if } (h,a)\in X^i \\ 0 & \text{if } (h,a)\notin X^i, p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))=0 \\ \frac{p(Y^i|\emptyset,\sigma)}{p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))} & \text{if } (h,a)\notin X^i, p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))\neq 0 \end{cases}$$

and

$$p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i})) = \sum_{\substack{h\in \hat{Y}^i,\exists h'\prec h, \\ h'\in X^i\neq Y^i}} p(h|\emptyset,(\sigma^i,\omega^{-i}))$$

and

$$f_{X^i,(\sigma^i,\omega^{-i})}(\cdot) = \left(1-\prod_{z'\in Z_{IS}(X^i,(\sigma^i,\omega^{-i}))} 1_{\{z'\}}(\cdot)\right) \sum_{z\in Z_{IS}(X^i,(\sigma^i,\omega^{-i}))} \frac{1_z(\cdot)\cdot z}{\sum_{\substack{X\in P'^i_{IS}(Y'),Y'\in\mathscr{Y}^i, \\ p_{X^i}(Y'|\emptyset,(\sigma^i,\omega^{-i}))>0}} 1_{Z_{IS}(Y',(\sigma^i,\omega^{-i}))}(z)}$$
$$+\left(\prod_{z'\in Z_{IS}(X^i,(\sigma^i,\omega^{-i}))} 1_{\{z'\}}(\cdot)\right)\cdot \Phi^i((\sigma^i,\omega^{-i})|X^i)$$

The equation 3.1 describes that, the player i acts at the information set X^i. Player i does not know which node in the information set he is currently located. He has the probability of $\mu(h|X^i,(\sigma^i,\omega^{-i}))$ to be at the history h at that moment and $\sigma^i_{X^i}$ to play action a. Then, by the actions of other players, the player i reaches Y^i, Y^i could be an information set in $\mathscr{X}^i$ or the set containing a terminal node. The probability of reaching Y^i from the node (h,a) is $\sum_{h'\in\hat{Y}^i,\ X^i\in P'^i_{IS}(Y^i)} p(h'|(h,a),\omega^{-i})$. The reason why we express it as the probability conditional on the other players' general strategies is that there is no history where the player i acts on the path from h to $h'\in\hat{Y}^i$. λ shows the player i's attitude to the expected payoffs of Y^i. The term $p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))$ is the probability of reaching Y^i through the information set X^i. When the probability is positive, we say the i-IS-path from X^i to Y^i is feasible. Thus, an i-IS-path from X to Y is **feasible**, if for any two information sets $X_1, X_2\in\mathscr{X}^i$ such that $X_1\in P'^i_{IS}(X_2)$, we have $p_{X_1}(X_2)>0$. Denote $Z_{IS}(X^i,(\sigma^i,\omega^{-i}))$ the set of terminal nodes satisfying there is at least one feasible IS-path from X to the set containing that terminal node with the general strategy profile (σ^i,ω^{-i}). If the strategy profile (σ^i, ω^{-i}) is completely mixed, we have $Z_{IS}(X^i, (\sigma^i, \omega^{-i})) = Z_{IS}(X^i)$. A strategy is completely mixed if the probability of each strategy in its support is positive. As we have stated above, a player's evaluation of the expected payoffs at information set is the same at any section of that information set. If $X^i \neq Y^i$, the probability that the player i reach Y^i from the action at information set X^i is $1/\lambda$. The explanation could be that player i at information set X^i as an outsider of Y_i and observe $1/\lambda$ part of Y^i and the evaluation of that part is $E^{IS}u(Y^i)$. Then, the evaluation of the expected payoffs at Y^i of the player i at information set X^i should be $E^{IS}u(Y^i)/(1/\lambda) = \lambda E^{IS}u(Y^i)$. Besides it, define the function $f_{X^i,(\sigma^i,\omega^{-i})}(\cdot) : \times_{|Z|}(\Re \times Z) \to X_{|Z|}(\Re \times Z)$ the **average function** which assigns equal probability to each feasible i-IS-path from the information set X^i to the set containing the same terminal node with the general strategy profile (σ^i, ω^{-i}). The reason why we call it average function is that the denominator

$$\sum_{\substack{X \in P'^i_{IS}(Y'), Y' \in \mathscr{Y}^i, \\ p_{X^i}(Y'|\emptyset,(\sigma^i,\omega^{-i}))>0}} 1_{Z_{IS}(Y',(\sigma^i,\omega^{-i}))}(z)$$

is the number of i-IS-paths from X^i to $\{z\}$. The function indicates that among several i-IS-path to the set containing a terminal node, there is at most one containing the path. The probability of the i-IS-path to the set containing a terminal node is weighted among several feasible IS-paths by the function f. Therefore, we finish the first step of calculating the attitude probability of i-IS-path from the current information set to. the set containing each terminal node.

In the second step, we substitute a terminal node (an outcome) in Φ with the payoffs of that terminal node by a **substitution function**. Define $g(\cdot) : \times_{|Z|}(\Re \times Z) \to \Re$ the substitution function, and

$$g^i(\cdot) = \sum_{z \in Z} u_i(z) 1_{\{z\}}(\cdot).$$

We call the calculation rule of the expected payoffs of a sophisticated player i the sophisticated recursive calculation rule since the calculation rule is formulated in the form of recursion. Then, we should calculate the expected payoff function following the calculation rule. We call the resulted expected payoff function **IS expected payoff function**[1], denoted by $E^{IS}u$. Then, the IS expected payoff function at the information set X^i

[1]The "IS" in the definition is the abbreviation of "Information Set". In our paper, a definition contains the abbreviation "IS" if it implies the viewpoint of a sophisticated player who considers games with imperfect recall over his information sets and other player's strategies. A sophisticated player realises he meets identical situation at any node in a non-singleton information set and would take the same action at each node in equilibrium. Thus, the sophisticated player focuses on things happened at an information set where

with the general strategy profile (σ^i, ω^{-i}) is

$$E^{IS}u_i(X^i|(\sigma^i, \omega^{-i})) = \frac{1}{p(X^i|\emptyset, (\sigma^i, \omega^{-i}))} \sum_{z \in Z_{IS}(X^i, (\sigma^i, \omega^{-i}))} p(z|\emptyset, (\sigma^i, \omega^{-i}))u_i(z).$$

If the strategy profile (σ^i, ω^{-i}) is completely mixed, then

$$E^{IS}u_i(X^i|(\sigma^i, \omega^{-i})) = \frac{1}{p(X^i|\emptyset, (\sigma^i, \omega^{-i}))} \sum_{z \in Z_{IS}(X^i)} p(z|\emptyset, (\sigma^i, \omega^{-i}))u_i(z).$$

Therefore, we have defined how should a player evaluates an information set. We should discuss the solution concepts based on the IS expected payoff function and compare them with those defined in previous works.

3.6.2 Discussion on Deviations at Multiple Information Sets

There are several reasons that we need to consider the potential deviations at multiple information sets, and thus, we should also define the IS expected payoff function at a set of information sets based on the sophisticated recursive calculation rule. Firstly, as we have stated above, there might be cross-branch information sets in the relation $\prec'$. In that case, a player does not know whether he has passed the other information sets which are cross-branch with the current one. Thus, he does not know whether he could change the behaviours at the other information sets. When the player reevaluate different strategy at the current information set, he must assume his behaviours at other cross-branch information sets are fixed. If he is allowed to reconsider simultaneously at all cross-branch information sets, there might be an alternative strategy which reaches higher expected payoffs but require the player to deviate at multiple information sets. Furthermore, in games with perfect recall, we are able to define the partial order $\prec$ over all information sets while, with the presence of imperfect recall, the partial order cannot be applied to all information sets. For example, see Figure 3.10, we cannot construct the partial order $\prec$ on $\{e_1\}$ and X. However, by allowing reconsideration at multiple information sets, we could discuss the partial order $\prec$ between $\{e_1\} \cup \{e_2\}$ and X, $\{e_1\} \cup \{e_2\} \prec X$. Thus, the properties satisfied under the relation $\prec$ are also satisfied in games with imperfect recall.

---

he acts rather than a particular node in a that information set. However, his viewpoint regarding other players is the same as if he does not present imperfect recall.

Assume that a player i evaluates the game at a set of information sets $\mathscr{X}_0^i \in \mathscr{X}^i$, then the sophisticated recursive calculation rule is

$$\Phi^i((\sigma^i,\omega_{-i})|\mathscr{X}_0^i) = f_{\mathscr{X}_0^i,(\sigma^i,\omega^{-i})}\Big(\sum_{h\in\mathscr{X}_0^i}\sum_{a\in A(h)}\mu(h|\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))\sigma_h^i(a)\sum_{\substack{h'\in\hat{Y}^i,\exists X^i\in\mathscr{X}_0^i,\\ s.t. X^i\in P_{IS}'^i(Y^i)}} p(h'|(h,a),\omega^{-i})\cdot \\ \lambda(h,a|(\sigma^i,\omega^{-i}),\mathscr{X}_0^i)\Phi^i((\sigma^i,\omega^{-i})|Y^i)\Big). \tag{3.2}$$

where

$$\lambda(h,a|(\sigma^i,\omega^{-i}),\mathscr{X}_0^i) := \begin{cases} 1 & \text{if } (h,a)\in X^i \\ 0 & \text{if } (h,a)\notin X^i, p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))=0 \\ \frac{p(Y^i|\emptyset,(\sigma^i,\omega^{-i}))}{p_{\mathscr{X}_0^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))} & \text{if } (h,a)\notin X^i, p_{X^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))\neq 0 \end{cases}$$

and

$$p_{\mathscr{X}_0^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i})) = \sum_{\substack{h\in\hat{Y}^i,\exists h'\prec h,\\ h'\in\mathscr{X}_0^i\neq Y^i}} p(h|\emptyset,(\sigma^i,\omega^{-i}))$$

and

$$f_{\mathscr{X}_0^i,(\sigma^i,\omega^{-i})}(\cdot) = \left(1-\prod_{z'\in Z_{IS}(\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))}1_{\{z'\}}(\cdot)\right)\sum_{z\in Z_{IS}(\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))}\frac{1_z(\cdot)\cdot z}{\sum_{\substack{\exists X^i\in\mathscr{X}_0^i\ s.t.\\ X^i\in P_{IS}'^i(Y'),\ Y'\in\mathscr{Y}^i,\\ p_{\mathscr{X}_0^i}(Y'|\emptyset,(\sigma^i,\omega^{-i}))>0}} 1_{Z_{IS}(Y',(\sigma^i,\omega^{-i}))}(z)} \\ +\left(\prod_{z'\in Z_{IS}(\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))}1_{\{z'\}}(\cdot)\right)\cdot\Phi^i((\sigma^i,\omega_{-i})|\mathscr{X}_0^i).$$

Now we define the multi-agent form of an extensive finite. In a multi-agent formation of the game, each agent is assigned to a set of information sets. The agent is able to control his behaviours at any information set in that set. Thus, the agent's control power is different when the multi-agent formation is different. At an information set X, if we model an agent plays at the set of information sets including X and every information set to which there is an IS-path from X. Then, the strategy which is conditionally optimal at X for a player belongs to the set of best responses given the strategy profile of other players and his strategy at other information set which is excluded from the set of information sets. Thus, if a player chooses an optimal behavioural rule at any set of information sets according to the sophisticated recursive calculation rule, then he chooses a strategy satisfies a stronger version of one-shot deviation principle. It requires the strategy to be optimal at the set of information sets to which there is an IS-path from the current information set.

Result 3.6.1 *For any player i, the IS expected payoff function from the sophisticated recursive calculation rule at the set of information sets* $\mathscr{X}_0^i$ *with a general strategy profile* (σ^i, ω^{-i}) *is*

$$E^{IS}u_i(\mathscr{X}_0^i|(\sigma^i,\omega^{-i})) = \frac{1}{p(\mathscr{X}_0^i|\emptyset,(\sigma^i,\omega^{-i}))} \sum_{z\in Z_{IS}(\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))} p(z|\emptyset,(\sigma^i,\omega^{-i}))u_i(z),$$

and its attitude modified IS-path probability function is

$$\Phi(\mathscr{X}_0^i|(\sigma^i,\omega^{-i})) = \frac{1}{p(\mathscr{X}_0^i|\emptyset,(\sigma^i,\omega^{-i}))} \sum_{z\in Z_{IS}(\mathscr{X}_0^i,(\sigma^i,\omega^{-i}))} p(z|\emptyset,(\sigma^i,\omega^{-i}))\cdot z$$

Proof of 3.6.1. See the Appendix. The IS expected payoff function at an information set X is a special case of result 1 if $\mathscr{X}_0^i = \{X\}$.

Although the belief system does not appear in the IS expected payoff function, the existence of a belief system μ consistent with the general strategy profile ω is required in the sophisticated recursive calculation rule. Thus, the implicit requirement for the IS expected payoff function is a belief system consistent with the strategy being evaluated conditional at the current information set. It is obviously that the IS expected payoff function coincides with the classic expected payoff function in games with perfect recall. In the situation of perfect recall but imperfect information, if there is a non-singleton information set for player i, he might not be informed of the moves of other players or Nature. At a non-singleton information set of player i, given the strategy profile σ and its consistent belief system μ, the expected payoff function is

$$\begin{aligned} Eu_i(\sigma|X) &= \sum_{h\in X}\mu(h|X,\sigma)\sum_{z\in Z}p(z|h,\sigma)u_i \\ &= \sum_{h\in X}\frac{p(h|\emptyset,\sigma)}{\sum_{h'\in X}p(h'|\emptyset,\sigma)}\sum_{z\in Z}p(z|h,\sigma)u_i(z) \\ &= \frac{1}{\sum_{h'\in X}p(h'|\emptyset,\sigma)}\sum_{z\in Z(X)}p(z|\emptyset,\sigma)u_i(z) \end{aligned}$$

In games with perfect recall, there is no absentmindedness in the game tree, thus

$$\frac{1}{\sum\limits_{h'\in X} p(h'|\emptyset,\sigma)} = \frac{1}{\sum\limits_{h'\in \hat{X}} p(h'|\emptyset,\sigma)} = \frac{1}{p(X|\sigma)}.$$

Additionally, $Z(X) = Z_{IS}(X)$ since the i-IS-path from the information set X to a set containing a terminal node is unique. It contains all the paths from any history in X to the terminal node. Therefore, the IS expected payoff function $E^{IS}u_i$ is the same as the classic expected payoff function Eu_i in games with perfect recall.

All players choose their ex-ante strategies before the game starts. If we interpret the ex-ante period as making decisions at $\mathscr{X}$, then the IS expected payoff function of player i at $\mathscr{X}$ with a general strategy profile ω is

$$E^{IS}u_i(\omega|\mathscr{X}) = \frac{1}{p(\mathscr{X}|\emptyset,\omega)} \sum_{z\in Z_{IS}(\mathscr{X},\omega)} p(z|\emptyset,\omega)u_i(z) = \sum_{z\in Z} p(z|\emptyset,\omega)u_i(z).$$

Thus, the $Eu_i(\omega|\emptyset) = E^{IS}u_i(\omega|\mathscr{X})$. The IS expected payoff function for the whole game is identical to the ex-ante expected payoff function.

3.7 IS Sequential Equilibrium and Different Solution Concepts

In this section, we are going to propose the solution concepts in terms of the IS expected payoff function and discuss the inclusion and existence of them. The solution concepts are defined in a similar way as those defined in Halpern & Pass (2016).

3.7.1 An Example

Before the further discussion on the solution concepts, we apply the IS expected payoff function to an example, compute the sequential equilibria of it and then compare them with the sequential equilibria in the classic expected payoff function. We borrow the game in Hillas & Kvasov (2020a), see Figure 3.13. The corresponding table of payoffs of player 1 and 2 is presented in Table 3.1. It is easy to see there are three pure strategy equilibria: $(T,LU),(T,RU)$ and (B,LD) in this game.

The Nash equilibrium (B,LD) is a sequential equilibrium without controversy. At any information set, both players do not have profitable deviations, no matter they are allowed to have one-step deviation or deviating to a totally different strategy. The Nash equilibrium (T,LU) is not a sequential equilibrium. If he is allowed to deviate at the current information set and regards the action at other information set as fixed, he can increase his payoff by deviating from the action L (payoff of 0) to R (payoff of 1). If he is allowed to deviate at both information sets, he can even deviate to LD (payoff of 3) which brings him higher expected payoffs than changing his current behaviour. In both cases, the player 2 has profitable deviations. Thus, (T,LU) is definitely not a sequential equilibrium. As for the Nash equilibrium (T,RU), the player 2 deviates to (L,D) if he could deviate at both information sets but had no profitable one-shot deviation. Thus, it satisfies the principle of sequential rationality only if he is allowed to deviate at a single information set. This example also illustrates a strategy satisfies no one deviation property might not be a sequential equilibrium.

However, when the players evaluate the game in terms of IS expected payoff function, both (B,LD) and (T,RU) are sequential equilibria not only in the situation that the player 2 can change his immediate action but also in the situation that he can switch to an alternative strategy. Denote the player 1's strategy is $(T,B)=(p,1-p)$ and the player 2's strategy is $(L,R;U,D)=(q_1,1-q_1;q_2,1-q_2)$. The player 2's IS expected payoff function at $\{e_1\}$ is

$$E^{IS}u_2(\{e_1\}|(p,q_1,q_2))=(1-q_1)+3q_1(1-q_2)\frac{2p}{1-p}q_2.$$

When the player 1 plays B ($p=0$), the best response of player 2 is $q_1=1,q_2=0$. When the player 1 plays T ($p=1$), the denominator in the term $2p/(1-p)$ is 0. In this case, we assume the player 2 reaches at $\{e_1\}$ because of player 1's small mistake. Then, assume $p=\varepsilon$ and ε is a very small number. The term $2p/(1-p)$ is extremely large. The player 2's conditional optimal strategy is $(q_1,q_2)=(0,1)$, i.e. (R,U). The player 2's IS expected payoff function at X is

$$E^{IS}u_2=\frac{3(1-p)q_1+[2p-3(1-p)q_1]q_2}{p+(1-p)q_1}.$$

When the player 1 plays B and the player 2 plays L ($(p,q_1) = (0,1)$), player 2's optimal strategy at X is $q_2 = 0$ (D). When the player 1 plays T and the player 2 plays R ($(p,q_1) = (1,0)$), the player 2' optimal strategy at X is $q_2 = 1$ (U).

3.7.2 Switch of a Strategy during Execution

When a player reconsiders in the middle of a game, he might deviate to a new strategy if the current one is not conditionally optimal. Several subtle issues should be illustrated regarding the switch of a strategy. In this paper, players are assumed to implement a general profile in games with imperfect recall. They choose randomisation over behavioural strategies before a game starts. Once the game plays, only one behavioural strategy from the support of the general strategy he has chosen ex-ante is actually played. If a player deviates to another strategy during execution, he should deviate to another behavioural strategy. Therefore, we should specify the deviation of a strategy in behavioural strategies. Now we define two types of deviations. Define $[\sigma^i, X, \sigma_0^i]$ a strategy of player i according to which i players σ_0^i at information set X and each information set X' such that $X \prec X'$, otherwise i plays the behavioural strategy σ^i. As stated in Halpern & Pass (2016), this type of deviations is mainly applied to the discussion over sequential rationality over the partial order $\prec$. In our framework, define $[\sigma^i; (\sigma'^i, X)]$ a strategy that the player i plays σ'^i at information set X and every information set X' at which player i moves, such that $X \prec' X'$ but not $X' \prec' X$, otherwise the player i plays σ^i. Notice that, the strategy $[\sigma^i; (\sigma'^i, X)]$ is a behavioural strategy if σ^i and σ'^i are both behavioural strategies since, for any information set $X^i \in \mathscr{X}^i$, the player i implements the same behavioural rule (either σ_X^i or $\sigma_X'^i$) at every history in X. Similarly, define $[\sigma^i; (\sigma'^i, \mathscr{X}_0^i)]$ a behavioural strategy that the player i plays σ'^i at information sets in $\mathscr{X}_0^i$ and information sets X' such that $\exists X'' \in \mathscr{X}_0^i, X'' \prec' X'$ and otherwise plays the behavioural strategy σ^i.

At the information set X in a game with perfect recall, a player considers a switch of strategy in the form of σ' instead of $[\sigma; (\sigma', X)]$ if the player's initial strategy is σ. In fact, the two strategies are indifferent to the player i since the player's conditionally optimal strategy is irrelevant to his previous actions. However, at a non-singleton information set with imperfect recall, the reason why the player i cannot differentiate histories might be he forgets his previous actions. The belief system he forms of the probability distribution on the histories could be determined by his own strategy, even his current behavioural rule on X. The belief system consistent with σ' might be different from that consistent with $[\sigma;(\sigma',X)]$ which makes a different expected payoff maximisation problem for player i. Thus, it is crucial to make clear what strategy he adopts at every information set X' such that $X' \prec' X$.

Now we compare the different forms of switching over different relationship among information sets. As explained in Halpern & Pass (2016), the behavioural strategy $[\sigma_i, X, \sigma_i']$ indicates that a player can change his behaviours at the information sets where he receives the same updated strategy at every history in that information set. We use the example shown in Figure 3.10 to compare different forms of switching. At $\{e_1\}$, the player cannot change his behavioural rule at X when he reevaluates the game by deviating to a strategy in the form of $[\sigma_i, X, \sigma_i']$. The explanation could be that the player cannot reach e_4 from e_1. Suppose there is a way to notice the switch of strategy when the player reaches the history where his behavioural rule at that history has been changed, the player's updated strategies at e_3 and e_4 are different, which violates the assumption that a player should be totally indistinguishable among histories in an information set. Another explanation could be that a player is not allowed to change the behaviours at the information sets where the switch might have effects on the execution of other incompatible paths. In Figure 3.10, if the player at X is informed of the updated strategy which has been switched at e_1, his behaviour at e_4 might be different, but the player at e_1 would never reach e_4. If the player changes his strategy at e_1 to the form σ', then he plays σ' at e_3 and plays σ at e_4 if ignoring the restriction on the information set. On the contrary, the strategy $[\sigma;(\sigma',X)]$ indicates the player can change the behavioural rules at the information sets which can be reached by IS-paths. We say there is an IS-path from information set X to X', the player can reach X' from X if the player is allowed to move among histories within an information set. In the example shown in Figure 3.10, the player at e_1 can change the behavioural rule at X if he switches to a strategy in the form of $[\sigma;(\sigma',X)]$. The assumption is consistent with our framework. In our framework, the player knows how he evaluates at an information set that comes later and considers it when he reconsiders at the current information set. Thus, the player knows he at a later information set will be informed of his current switch and cannot differentiate the histories in that information set. Then, he evaluates the game as if he plays the switched strategy at each history, although he cannot play the switched strategy at a node he cannot reach. Under this assumption, it is not problematic to allow the player to change his behaviours at an information set in which not all histories can be reached from the player's current history.

3.7.3 Perturbed Games

The sequential rationality requires the strategy of each player is conditionally optimal at all information sets given their belief systems. We call those information sets which are reached with positive probability by a Nash equilibrium compose the on-equilibrium path and the rest information sets in $\mathscr{H}$ belong to the off-equilibrium path. For any Nash equilibrium in general strategies ω, and any information set X where $p(X|\emptyset,\omega) > 0$, a reasonable belief system on X should be the one derived from the strategy profile ω through Bayes' rule. However, how to define beliefs at information sets in off-equilibrium path becomes an issue. We should solve it before officially define sequential equilibrium in terms of IS expected payoff function.

In an extensive game with imperfect recall, denoted by Γ, for any information set X, define a function $\eta : A(X) \to \Re([0,1])$ which assign a probability $\eta_a \in (0,1)$ to each available action $a \in A(X)$ such that $\sum_{a \in A(X)} \eta_a < 1$. We call η a perturbation of Γ. The η_a is interpreted as a small tremble or mistake of players. Every η_a is a very small probability since mistakes are very unlikely. We say a sequence of functions $\eta_n \to 0, n \to \infty$ if $\eta_n(a) \to 0$ for every $a \in A$.

Define (Γ,η) a **perturbed game** which is composed by a game Γ and a perturbation η. For any player i, we say the behavioural strategy σ^i is acceptable if, for any action $a \in A(X^i)$, $\sigma^i_X(a) \geq \eta_a$. A general strategy ω^i is acceptable if, any behavioural strategy from the support of ω^i is acceptable. Notice that, if a general strategy profile ω is acceptable in a perturbed game (Γ,η), then for any information set X^i where the player i moves, $p(X^i|(\emptyset,\omega)) > 0$ and $p(X^i|\emptyset,(\sigma^i,\omega^{-i})) > 0$ for any behavioural strategy σ^i from the support of his general strategy ω^i. Therefore, we could apply Bayes' rule to any information set to form a reasonable belief system. For any player i, a behavioural strategy σ^i is a completely mixed behavioural strategy if it is acceptable in a perturbed game (Γ,η) and a general strategy ω^i is a completely mixed general strategy if it is acceptable in a a perturbed game (Γ,η). Any general strategy profile referred in the solution concepts are assumed to be acceptable.

3.7.4 Perfect Equilibrium

According to Selten (2020), in a finite normal form game with perfect recall, a mixed strategy profile π is a trembling hand perfect equilibrium if there exists a sequence of completely mixed strategy profiles π_k which converges to π such that, for each player i, his strategy π^i is a best response to π_k^{-i} for any k. It describes that, for any player i, his equilibrium strategy π^i is always the optimal strategy if the other players make small mistakes when they play their equilibrium strategy profile π^{-i}. Then, a perfect equilibrium is robust with small perturbations of other players. There are two ways to extend the definition of perfect equilibrium to the games in extensive forms. The first one is to apply the above condition to extensive games. A Nash equilibrium satisfying the above condition is called a normal form trembling hand perfect equilibrium. In the other way of extension, a Nash equilibrium π in mixed strategies is an extensive form trembling hand perfect equilibrium (perfect equilibrium) in a extensive game Γ if there is a sequence of perturbed games (Γ, η_k), and a sequence of completely mixed strategy profiles π_k such that $\eta_k \to 0$, π_k is a Nash equilibrium of (Γ, η_k), and π_k converges to π.

In this paper, we restrict the attention to extensive form trembling hand equilibria (perfect equilibria). In the presence of perfect recall, perfect equilibria are defined over mixed strategies. Correspondingly, perfect equilibria should be defined over general strategies in extensive finite games with imperfect recall.

Definition 3.7.1 *In an extensive finite game with imperfect recall Γ, a general strategy profile ω is an* ***perfect equilibrium****, if there is a sequence of perturbed games (Γ, η_k) and a sequence of completely general strategy profiles ω_k satisfy that, $\eta_k \to 0, k \to \infty$, ω_k is a Nash equilibrium of (Γ, η_k) and ω_k converges to ω.*

A perfect equilibrium in general strategies exists in every finite game. Before officially proving the existence of such a perfect equilibrium, we cite the definition of Kuhn-equivalence and express a proposition regarding it.

Definition 3.7.2 *Hillas et al. (2020b) Two strategies ω_1^i, ω_2^i of player i are said to be* ***Kuhn-equivalent*** *if, for any general strategy profile $\omega^{-i} \in \Omega^{-i}$ of other player, the strategy profiles $(\omega_1^i, \omega^{-i})$ and $(\omega_2^i, \omega^{-i})$ induce the same probability distribution over the the terminal nodes.*

Proposition 3.7.1 *(Isbell 1957, 1959, Alpern 1988, Halpern & Pass 2016, Hillas et al. 2020b) For any player i in a finite game Γ, there is a finite number K_Γ which only depends on Γ, such that for any general strategy ω^i of player i, there is a general strategy ω'^i assigning positive probability to at most K_Γ elements from Σ^i that is Kuhn-equivalent to ω^i.*

The proposition 3.7.1 describes that any general strategy ω^i of player i has a Kuhn-equivalent general strategy ω'^i whose support has at most K_Γ elements. Then, it makes the set of completely general strategies satisfy the conditions in Kakutani's fixed point theorem since the set of acceptable general strategies is a compact and convex subset in $\Re^{K_\Gamma}$ if we rewrite ω'^i of player i as the a tuple of probabilities on each behavioural strategy in its support. The existence of Nash equilibria is invariable among Kuhn-equivalent strategy profiles since both Nash equilibria and Kuhn-equivalence are discussed over the probability measures over terminal nodes.

Theorem 3.7.1 *A perfect equilibrium exists in all finite games.*

Proof of Theorem 3.7.1 For an extensive finite game Γ, construct a sequence of perturbed games (Γ, η_k) such that $\eta_k \to 0, k \to \infty$. By Kakutani's fixed point theorem, there exists a Nash equilibrium in each perturbed game, denoted by ω_k. We have explained above that the set of acceptable completely mixed strategies is a compact subset C of $\Re$. Every sequence in C has a convergent subsequence and the limit point is also in C. Denote ω a limit point of $\omega_k, k \to \infty$.Then, ω is a perfect equilibrium in Γ. Denote μ_k the weakly consistent belief system with ω_k. Then, there exists a belief system $\mu = \lim_{k\to\infty} \mu_k$, such that μ is weakly consistent with ω.

Proposition 3.7.2 *In an extensive finite game Γ, a general strategy profile ω is a perfect equilibrium. For any player i, any behavioural strategy σ^i from the support of ω^i, there exists a belief system μ_{σ^i} consistent with*

(σ^i, ω^{-i}) *such that, for any set of information sets* $\mathscr{X}_0^i \subseteq \mathscr{X}^i$ *that the player i moves at those information sets, and for any behavioural strategy* σ'^i*, we have*

$$E^{IS}u_i((\sigma^i, \omega^{-i})|\mathscr{X}_0^i) \geq E^{IS}u_i(([\sigma^i;(\sigma'^i, \mathscr{X}_0^i)], \omega^{-i})|\mathscr{X}_0^i). \tag{3.3}$$

Proof of Proposition 3.7.2 In the proof of Theorem 3.7.1, we know there exists a sequence of perturbed games (Γ, η_k) such that there exists a sequence of general strategition y profiles ω_k which are the Nash equilibrium for each perturbed game respectively and their consistent belief systems μ_k. Then, for any player i and any behavioural strategy σ^i from the support of ω^i, we have a subsequence of behavioural strategies σ_k^i which converges to σ^i. (We could choose a subsequence of σ_k^i which is a behavioural strategy in the support of the convergent subsequence of ω_k^i.) In the meanwhile, we also have the sequence of belief systems $\mu_{\sigma_k^i}$ consistent with $(\sigma_k^i, \omega_k^{-i})$ converges to μ_{σ^i}. Notice that we can always chooses a convergent subsequence from a subsequence, since in a compact subset of $\Re^n$, n is a finite positive integer, there always exists a convergent subsequence of any sequence. Thus, such a sequence which satisfies $\lim_{k\to\infty}(\sigma_k^i, \omega_k^{-i}) = (\sigma^i, \omega^{-i})$ and $\lim_{k\to\infty}\mu_{\sigma_k^i} = \mu_{\sigma^i}$.

Then we should prove that the perfect equilibrium ω satisfies the inequation 3.3. See the Appendix.

So far, we have proved the existence of a perfect equilibrium in all finite games and a perfect equilibrium satisfies sequential rationality in terms of IS expected payoff function. In the next subsection, we are going to define sequential equilibria in games with imperfect recall.

3.7.5 IS Sequential Equilibrium

In the previous section, we have compared three different forms of the switch of a strategy: $[\sigma, X, \sigma']$, σ' and $[\sigma;(\sigma', X)]$. The form of $[\sigma, X, \sigma']$ puts the most restrictions on the information sets where a player is allowed to change the behavioural rule while the form of $[\sigma;(\sigma', X)]$ puts fewest restrictions. On the contrary, the requirement for a strategy profile to satisfy sequential rationality is most strict in the switch form of $[\sigma;(\sigma', X)]$

over the relation $\prec'$ while it is least strict in the switch form of $[\sigma, X, \sigma']$ over the partial order $\prec$. In the example shown in Figure 3.10, when the switch form is $[\sigma;(\sigma', X)]$ over the relation $\prec'$, a strategy profile satisfies sequential rationality if it is conditionally optimal at e_1, e_2 and X based on the fact that the player is allowed to change the behavioural rule at his current history and X if he is currently at e_1 or e_2. When the switch form is $[\sigma, X, \sigma']$ over the partial order $\prec$, a strategy profile satisfies sequential rationality if it is conditionally optimal at e_1, e_2 and X based on the fact that the player is allowed to change his immediate behaviour since there is no information set X' satisfying $\{e_1\} \prec X'$ or $\{e_2\} \prec X'$.

Definition 3.7.3 *In an extensive finite game Γ, a pair $(\omega, \{\mu_{\sigma^i} : \sigma^i$ is in the support of $\omega^i, i \in N\})$ composed by a general strategy profile ω and a family of belief systems μ_{σ^i}, one for each player i and each behavioural strategy σ^i in the support of ω^i, is called an assessment family. The assessment family $(\omega, \{\mu_{\sigma^i} : \sigma^i$ is in the support of $\omega^i, i \in N\})$ is an* ***IS sequential equilibrium*** *if each belief system μ_{σ^i} is consistent with (σ^i, ω^{-i}), and for any player i and any set of information sets $\mathscr{X}_0^i \subseteq \mathscr{X}^i$, we have*

$$E^{IS}u_i((\sigma^i, \omega^{-i})|\mathscr{X}_0^i) \geq E^{IS}u_i(([\sigma^i;(\sigma'^i, \mathscr{X}_0^i)], \omega^{-i}))|\mathscr{X}_0^i),$$

for any behavioural strategy σ'^i.

Theorem 3.7.2 *An IS sequential equilibrium exists in all finite games.*

Proof. See Theorem 3.7.1 and Definition 3.7.1.

The definition of a sequential equilibrium requires a general strategy profile is conditionally optimal at any set of information sets where the same player moves that those information sets. It impose a more strict criterion for an assessment family to be a sequential equilibrium than the definition on each information set. Not only the conditional optimality needs to be examined at more combinations of information sets but also does it allow a more reasonable criterion of conditional optimality at cross-branch information sets and the information sets, such as $\{e_1\}$ in Figure 3.10, where the player might not reach every history in his immediate successor information sets.

The existence of an IS sequential equilibrium indicates that there exists an equilibrium strategy profile such that the player with imperfect recall do not have motivation to change his current strategy at any information set, given the equilibrium strategy profile of the other players, if we use IS expected payoff function as the standard to examine the sequential rationality of a strategy. If the player with imperfect recall knows the fact of future imperfect recall and prepares for it, there exists a strategy that he thinks is conditionally optimal when he reconsiders at any information set. The way that the player prepares is to take how he will evaluate at the future information set with imperfect recall into his current consideration.

If we replace a strategy in a Nash equilibrium with an outcome-equivalent strategy, then the new strategy profile is also a Nash equilibrium. In games with perfect recall, for any general strategy, there exist an outcome-equivalent strategy in behavioural strategies and mixed strategies respectively. Thus, an assessment family is a sequential equilibrium in terms of IS expected payoff function if and only if an assessment consisting of an outcome-equivalent behavioural strategy and its consistent belief is a sequential equilibrium.

The expected payoff function of player i in Halpern & Pass (2016) could be written as

$$Eu(\omega|X) = \frac{1}{p(X|\emptyset,\omega)} \sum_{z \in Z(X)} p(z|\emptyset,\omega) u_i(z),$$

and

$$Eu(\omega|\mathscr{X}_0) = \frac{1}{p(X|\emptyset,\omega)} \sum_{z \in Z(\mathscr{X}_0)} p(z|\emptyset,\omega) u_i(z).$$

It is easy to see that the IS expected payoff function is the same as this expected payoff function wherever the general strategy profile satisfies $Z(X) = Z_{IS}(X,\omega)$,and $Z(\mathscr{X}_0) = Z_{IS}(\mathscr{X}_0,\omega)$.

3.7.6 IS Nash Equilibrium, Nash Equilibrium and Agent Equilibrium

We define the IS Nash equilibrium, Nash equilibrium and agent equilibrium in general strategies in this subsection and discuss the equilibrium hierarchy among perfect, sequential, Nash and agent equilibrium.

Definition 3.7.4 *In an extensive finite game Γ, a general strategy profile ω is a **IS Nash equilibrium** if and only if there exists a family of belief systems μ_{σ^i}, one for each player i and each behavioural strategy σ^i in the support of ω^i such that each belief system μ_{σ^i} is consistent with (σ^i, ω^{-i}),and for any player i and any set of information sets $\mathscr{X}_0^i \subseteq \mathscr{X}^i$ such that $p(\mathscr{X}_0^i|\emptyset, (\sigma^i, \omega^{-i})) > 0$, we have*

$$E^{IS}u_i((\sigma^i, \omega^{-i})|\mathscr{X}_0^i) \geq E^{IS}u_i(([\sigma^i; (\sigma'^i, \mathscr{X}_0^i)], \omega^{-i}))|\mathscr{X}_0^i),$$

for any behavioural strategy σ'^i.

We can see that the difference between the definition of an IS sequential equilibrium and an IS Nash equilibrium is that IS Nash equilibrium only requires the sequential rationality at the collection of information sets which could be reached with positive probability. Thus, an IS sequential equilibrium must be an IS Nash equilibrium.

Definition 3.7.5 *In a finite game Γ, we say a general strategy profile ω is a **Nash equilibrium**, if for any player i, and any behavioural strategy σ^i in the support of ω'^i, we have*

$$Eu_i(\sigma^i, \omega^{-i}|\emptyset) \geq Eu_i((\sigma'^i, \omega^{-i})|\emptyset).$$

for any behavioural strategy σ'^i.

We have stated in previous section that the IS expected payoff function is identical to ex-ante expected payoff function if assuming the set of information sets $\mathscr{X}_0^i = \mathscr{X}$. Thus, an IS Nash equilibrium is a Nash equilibrium. Although the existence of an IS Nash equilibrium shows the existence of a Nash equilibrium, we can also prove the existence by Kakutani's fixed point theorem. The proof is identical to that of the existence of a perfect equilibrium.

An agent equilibrium of the extensive game Γ is a strategy profile which is a Nash equilibrium of the agent form of Γ. The agent form of Γ is phrased as following: Every player in the extensive game is divided into several agents, each at one of the information sets where he moves, the utility function of all agents of a player is identical to that of the player.

Definition 3.7.6 *In a finite game Γ, we say a behavioural strategy profile σ is an* ***agent equilibrium****, if for any player i, any information set X^i, we have*

$$E^{IS}u_i(\sigma|\emptyset) \geq E^{IS}u_i((\sigma_X^{\prime i}, \sigma_{-X}^i, \sigma^{-i})|\emptyset),$$

for any $\sigma_X^{\prime i} \in \Sigma_X^i$.

The agent form models Γ as a simultaneous-move game where all the agents move simultaneously and maximises their expected payoff given the behavioural rules of other agents. Thus, the utility function of an agent is the ex-ante expected payoff of the player that the agent belongs to. The agent equilibrium is defined over behavioural strategies. A strategy profile is an agent equilibrium if the player cannot increase his ex-ante expected payoff by changing his behavioural rule at any information set. Although the agent form also analyses one-information set deviations, it is different from our approach. The agent form discuss one information set deviation before the game plays, while our approach examines one information set deviation based on the fact of choosing the same deviation at each history of an information set. Our framework examines the one information set deviation during execution. In games with perfect recall, the two approaches are equivalent. A general strategy profile is an agent equilibrium if every outcome-equivalent behavioural strategy profile is an agent equilibrium.

The existence of an agent equilibrium can also be proved by Kakutani's fixed point theorem in the same way as the proof of the existence of a perfect equilibrium. It is easy to see a strategy profile is an agent equilibrium if it is a Nash equilibrium. Given the equilibrium strategy profile of other players, if a strategy belongs to a

Nash equilibrium, the player who implements that strategy cannot increase his expected payoffs by changing his behaviour rules at any set of information sets where he moves. If a strategy belongs to an agent form, the player who implements that strategy cannot increase his expected payoffs by changing his behavioural rule at any single information set. A strategy is optimal with respect to deviations at any combination of multiple information sets must be optimal with respect to deviation at any single information set.

Theorem 3.7.3 *In all extensive finite games (perfect recall, linear, nonlinear), the assessment consisting of every perfect equilibrium and its consistent belief system is an IS sequential equilibrium, the strategy profile in every IS sequential equilibrium is an IS Nash equilibrium, every IS Nash equilibrium is a Nash equilibrium, and every Nash equilibrium is an agent equilibrium*

Therefore, by defining a sequential equilibrium and a Nash equilibrium based on the IS expected payoff functions, the chain of inclusion among different types of equilibrium in games with perfect recall is valid in all finite games.

If we use IS expected payoff function as the standard to examine sequential rationality at an information set, then Theorem 3.7.1 indicates that there exists a perfect equilibrium strategy profile such that, for any player, given the strategy profile of other players, the player will not be motivated to deviate any an information set, even if all the players have shaking hands and make small mistakes that they play the behavioural strategy which is not the support of the equilibrium strategy with a small probability. For each player, the small mistakes of the other players cannot motivate him to deviate from the current strategy at any information set he moves. The perfect equilibrium strategy profile is locally stable in terms of an unexpected behavioural strategy played by another player with a small probability.

Theorem 3.7.3 indicates that, with the assumption that a sophisticated player use IS expected payoff function as his standard to examine sequential rationality at any collection of information sets, a strategy profile is sequentially rational at any collection of information sets and is locally stable, then the strategy profile must be sequentially rational at any collection of information sets if there exists a family of belief systems which is consistent with this strategy profile. Then, a strategy profile must be sequentially rational at any collection of information sets such that the probability of reaching one of the collections of information sets is positive if there exists a family of belief systems which is consistent with this strategy profile and the strategy profile is sequentially rational at any collection of information sets.

To compare the difference between our definitions and the definitions in conventional game theory. There are mainly three aspects. Firstly, the available strategy in our definitions is a general strategy rather than a behavioural strategy or a mixed strategy. There exist games such that a behavioural strategy obtains more expected payoffs than any mixed strategies and there exist games such that a behavioural strategy obtains more expected payoffs than any mixed strategies. When a sophisticated player evaluates the game in the midst of execution of a game, he could choose any behavioural strategy, and compares the expected payoffs of the behavioural strategy to his current behavioural strategy, conditional on using the IS expected payoff function and the general strategy profile of other players. For this player, he has known the behavioural strategy he is implementing once the game starts. However, his information about other players' strategies is still the general strategy profile except for the situation that he reaches some information set which should not be reached if the other players implement the ex-ante chosen general strategy profile. Additionally, the requirement for consistent belief system is extended to a family of belief systems. For each behavioural strategy in the support of a general strategy, there should be a belief system consistent with the behavioural strategy. The second difference is that we examine the sequential rationality over collections of information sets instead of each information set. It is due to the complicated structures of games with imperfect recall. A strategy profile is sequentially rational at each information set does not represent it is also sequentially rational at any collection of information sets. The most salient difference is that in our definitions, the standard to examine sequential rationality is IS expected payoff function rather than the conventional conditional expected payoff function. It represents how a sophisticated player defines what is sequential rationality.

3.8 Conclusions and Discussions

In all finite games, players are assumed to know the structure of the game tree. Thus, a player who has imperfect recall knows the fact before and after the game starts. We propose that, if a player with imperfect recall is sophisticated, he should take the fact that he will forget the previous actions and he does not know which node in the information set he is currently located due to his own limited memory at the later stage of the game, into consideration when he evaluates whether his current strategy is conditionally optimal at an earlier stage. When a player reevaluates at an information set with imperfect recall, he evaluates the expected payoffs at every history and forms his subjective belief on the probability of currently being at every history. His expected payoff at that information set should be the sum of the expected payoff at each history weighted by the subjective probability of that history. Thus, his evaluation of the information set is the same no matter which history he is at physically. When he reconsiders at an information set which is an immediate predecessor information set of the information set with imperfect recall, he knows he will forget where he is and evaluate the information set in the way we state above. Then, at the predecessor information set, it is reasonable for the player to consider those terminal nodes which he knows he will never reach at the current information set if he notices the possibility of reconsideration at the information set with imperfect recall. In games with perfect recall, it does not matter whether a player is sophisticated.

We develop the sophisticated recursive calculation rule in all finite games and compute the IS expected payoff function by the calculation rule. IS expected payoff function provides us with the criterion for sophisticated players to examine the conditional optimality of a behavioural strategy at any information set based on the equilibrium strategy profile of other players. Then, we define an IS sequential equilibrium based on the assessment family and prove the existence of such an equilibrium in all finite games.

A general strategy profile satisfies a stronger version of one-shot deviation principle if it is an IS sequential equilibrium. A general strategy in an IS sequential equilibrium should be optimal conditionally on multiple formations of a multi-agent form of the game. The IS sequential equilibrium imposes players' best responses with the restriction of deviations at one or multiple information sets. The control power assigned to a player in the definition of IS sequential equilibrium is between the sequential equilibrium proposed by Halpern & Pass (2016), which allows players to deviate to another strategy; and the multiselves perfect equilibrium proposed by Lambert et al. (2019), which allows players change their behaviours at the current history.

Although we have discussed several problems arising due to imperfect recall, numbers of issues have not been solved or even mentioned, especially in a game presenting absentmindedness. For example, what could be a reasonable belief at a non-singleton information set with imperfect recall. Although the expected payoff at an information set seems to be irrelevant with his behavioural rule at the information set where he evaluates from the expression of IS expected payoff function, the implicit requirement is that the player's belief over histories in the information set might be determined by his behaviours at that information set. In our framework, we form a belief system conditional on any behavioural rule at that information set, and the player who moves at that information set chooses one in the set of best responses, and then requires the chosen optimal behavioural rule to be identical to the one which works as the evidence to form the belief system. It is difficult to explore the equilibrium solution concepts in games with imperfect recall. The reason behind it is that the equilibrium existence and equilibrium hierarchy could be totally different when we impose different assumptions on the players in a game with imperfect recall; while it results in the same equilibrium in spite of different assumptions of players in the presence of perfect recall. Besides technical issues, the intuitive explanation of some models is not reasonable or realised. For example, as we mentioned in previous sections, how could a player with imperfect recall always remember the updated strategy but forgets his previous actions? If players cannot remember the latest strategy, can we model games with imperfect recall in one-self approach? If not, is the multiself approach the unique way to model games with imperfect recall? It seems that we are still at the very beginning stage of modelling games with perfect recall.

3.9 Appendix

Proof. of Theorem 3.6.1.

For simplicity in notation, assume that the general strategy profile (σ^i, ω^{-i}) is completely mixed and rewrite $f_{\mathscr{X}^i_0,(\sigma^i,\omega^{-i})}$ as f and $(\sigma^i, \omega^{-i}) = \omega_i$.

Firstly, we compute a special case. For any player i, there must be an information set X^i or a set of information sets $\mathscr{X}_0^i \subseteq \mathscr{X}^i$ satisfying, for any $Y^i \in \mathscr{Y}^i$, such that $X^i \in P_{IS}'^i(Y^i)$ or $\mathscr{X}_0^i \in P_{IS}'^i(Y^i)$, we have $\exists z \in Z, Y^i = \{z\}$. For a more general case, we compute the IS expected payoffs at $\mathscr{X}_0^i$. For any $Y^i = \{z\}$, $\Phi^i(\omega_i|Y^i) = \Phi^i(\{z\}) = z$. The feasible IS-path number function is $f = \sum_{z \in Z} 1_{\{z\}}(\cdot) \cdot z$. The attitude modified IS-path probability

$$\begin{aligned}
\Phi^i(\omega_i|\mathscr{X}_0^i) &= f\left(\sum_{h \in \mathscr{X}_0^i} \sum_{a \in A(h)} \mu(h|\mathscr{X}_0^i, \omega_i)\sigma_h^i(a) \sum_{\substack{h' \in \hat{Y}^i, \exists X^i \in \mathscr{X}_0^i, \\ s.t. X^i \in P_{IS}'^i(Y^i)}} p(h'|(h,a), \omega^{-i}) \cdot \Phi^i(\omega_i|Y^i)\right) \\
&= \sum_{h \in \mathscr{X}_0^i} \sum_{a \in A(h)} \frac{p(h|\emptyset, \omega_i)\sigma_h^i(a)}{\sum_{h'' \in \mathscr{X}_0^i} p(h''|\emptyset, \omega_i)} \cdot \Phi^i(\omega_i|Y^i) \\
&= \frac{\sum_{h \in \mathscr{X}_0^i \setminus \widehat{\mathscr{X}_0^i}} p(h|\emptyset, \omega_i)}{\sum_{h'' \in \mathscr{X}_0^i} p(h''|\emptyset, \omega_i)} \Phi^i(\omega_i|\mathscr{X}_0^i) + \sum_{z \in Z(\mathscr{X})_0^i} p(z|\emptyset, \omega_i) \cdot z
\end{aligned}$$

where

$$f(\cdot) = \sum_{z \in Z(\mathscr{X}_0^i)} 1_{\{z\}}(\cdot) \cdot z,$$

and define $h \in \widehat{\mathscr{X}_0^i}$, if there is no $h' \in H$ satisfying $h' \prec h$ and $\exists X^i \in \mathscr{X}_0^i$, such that $h' \in X^i$. Then,

$$\left(\sum_{h'' \in \widehat{\mathscr{X}_0^i}} p(h''|\emptyset, \omega_i)\right) \Phi^i(\omega_i|\mathscr{X}_0^i) = \sum_{z \in Z(\mathscr{X}_0^i)} p(z|\emptyset, \omega_i) \cdot z.$$

Thus,

$$\Phi^i(\omega_i|\mathscr{X}_0^i) = \frac{1}{\sum_{h'' \in \widehat{\mathscr{X}_0^i}} p(h''|\emptyset, \omega_i)} \sum_{z \in Z(\mathscr{X}_0^i)} p(z|\emptyset, \omega_i) \cdot z$$

The corresponding IS expected payoff function is

$$E^{IS} u_i(\omega_i|\mathscr{X}_0^i) = \frac{1}{\sum_{h'' \in \widehat{\mathscr{X}_0^i}} p(h''|\emptyset, \omega_i)} \sum_{z \in Z(\mathscr{X}_0^i)} p(z|\emptyset, \omega_i) u_i(z).$$

We have compute the IS expected payoffs of a special case and its result satisfies the general term in result 1. Because of the complication to directly compute the general term of IS expected payoffs, we would rather verify whether the general term showed in result 1 satisfies the sophisticated recursive calculation rule.

$$\Phi^i(\omega_i|\mathscr{X}_0^i) = f\left(\sum_{h\in\mathscr{X}_0^i}\sum_{a\in A(h)}\mu(h|\mathscr{X}_0^i,\omega_i)\sigma_h^i(a)\sum_{\substack{h'\in\hat{Y}^i,\exists X^i\in\mathscr{X}_0^i,\\ s.t.X^i\in P_{IS}'^i(Y^i)}} p(h'|(h,a),\omega^{-i})\lambda(h,a|(\omega_i,\mathscr{X}_0^i)\Phi^i(\omega_i|Y^i)\right)$$

$$= f\left(\sum_{h\in\mathscr{X}_0^i}\sum_{a\in A(h)}\frac{p(h|\emptyset,\omega_i)\sigma_h^i(a)}{\sum_{h''\in\mathscr{X}_0^i}p(h''|\emptyset,\omega_i)}\sum_{\substack{h'\in\hat{Y}^i,\exists X^i\in\mathscr{X}_0^i,\\ s.t.X^i\in P_{IS}'^i(Y^i)}} p(h'|(h,a),\omega^{-i})\lambda(h,a|(\omega_i,\mathscr{X}_0^i)\Phi^i(\omega_i|Y^i)\right)$$

Then,

$$\sum_{h''\in\mathscr{X}_0^i}p(h''|\emptyset,\omega_i)\Phi^i(\omega_i|\mathscr{X}_0^i) = \sum_{h\in\mathscr{X}_0^i\setminus\widehat{\mathscr{X}_0^i}}p(h|\emptyset,\omega_i)\Phi^i(\omega_i|\mathscr{X}_0^i)$$

$$+f\left(\sum_{\substack{h'\in\hat{Y}^i,\exists X^i\in\mathscr{X}_0^i,\\ s.t.X^i\in P_{IS}'^i(Y^i)}} p(h'|\emptyset,\omega_i)\frac{p(Y^i|\emptyset,\omega_i)}{p_{\mathscr{X}_0^i}(Y^i|\emptyset,(\sigma^i,\omega^{-i}))}\frac{1}{p(Y^i|\emptyset,\omega_i)}\sum_{z\in Z_{IS}(Y^i)}p(z|\emptyset,\omega_i)\cdot z\right)$$

$$\sum_{h''\in\widehat{\mathscr{X}_0^i}}p(h''|\emptyset,\omega_i)\Phi^i(\omega_i|\mathscr{X}_0^i) = f\left(\sum_{\substack{\exists X^i\in\mathscr{X}_0^i,\ s.t.\\ X^i\in P_{IS}'^i(Y^i)}}\sum_{z\in Z_{IS}(Y^i)}p(z|\emptyset,\omega_i)\cdot z\right)$$

Thus,

$$\Phi^i(\omega_i|\mathscr{X}_0^i) = \frac{1}{\sum_{h''\in\widehat{\mathscr{X}_0^i}}p(h''|\emptyset,\omega_i)}\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_i)\cdot z$$

$$E^{IS}u_i(\omega_i|\mathscr{X}_0^i) = \frac{1}{\sum_{h''\in\widehat{\mathscr{X}_0^i}}p(h''|\emptyset,\omega_i)}\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_i)\cdot u_i(z),$$

since

$$f_{\mathscr{X}_0^i,\omega_i}(\cdot)=\left(1-\prod_{z'\in Z_{IS}(\mathscr{X}_0^i)}1_{\{z'\}}(\cdot)\right)\sum_{z\in Z_{IS}(\mathscr{X}_0^i,\omega_i)}\frac{1_z(\cdot)\cdot z}{\sum_{\substack{\exists X^i\in\mathscr{X}_0^i\ s.t.\\ X^i\in P_{IS}'^i(Y'),\ Y'\in\mathscr{Y}^i,\\ p_{\mathscr{X}_0^i}(Y'|\emptyset,\omega_i)>0}}1_{Z_{IS}(Y')}(z)}$$
$$+\left(\prod_{z'\in Z_{IS}(\mathscr{X}_0^i)}1_{\{z'\}}(\cdot)\right)\cdot\Phi^i(\omega_i|\mathscr{X}_0^i).$$

In conclusion, the general term of $E^{IS}u_i$ and Φ satisfy the sophisticated recursive calculation rule.

Proof. of Proposition 3.7.2.

Now we prove a perfect equilibrium ω must satisfies the inequation (3). For simplicity in notation, assume $\omega_1=(\sigma^i,\omega^{-i})$ and $\omega_2=([\sigma^i;(\sigma'^i,\mathscr{X}_0^i)],\omega^{-i})$.

Assume there exists a player i, a behavioural strategy σ'^i and a set of information sets $\mathscr{X}_0^i$ such that

$$E^{IS}u_i(\omega_1|\mathscr{X}_0^i)<E^{IS}u_i(\omega_2|\mathscr{X}_0^i).$$

Then,

$$\begin{aligned}
\frac{1}{p(\mathscr{X}_0^i|\emptyset,\omega_1)}\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_1)u_i(z)&<\frac{1}{p(\mathscr{X}_0^i|\emptyset,\omega_2)}\sum_{z\in Z_{IS}(\mathscr{X}_0^i,\omega_2)}p(z|\emptyset,\omega_2)u_i(z)\\
\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_1)u_i(z)&<\sum_{z\in Z_{IS}(\mathscr{X}_0^i,\omega_2)}p(z|\emptyset,\omega_2)u_i(z)\\
\sum_{z\in Z\setminus Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_1)u_i(z)+\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_1)u_i(z)&<\sum_{z\in Z\setminus Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_2)u_i(z)+\sum_{z\in Z_{IS}(\mathscr{X}_0^i,\omega_2)}p(z|\emptyset,\omega_2)u_i(z)\\
\sum_{z\in Z}p(z|\emptyset,\omega_1)u_i(z)&<\sum_{z\in Z\setminus Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_2)u_i(z)+\sum_{z\in Z_{IS}(\mathscr{X}_0^i)}p(z|\emptyset,\omega_2)u_i(z)\\
\sum_{z\in Z}p(z|\emptyset,\omega_1)u_i(z)&<\sum_{z\in Z}p(z|\emptyset,\omega_2)u_i(z)\\
Eu_i(\omega_1|\emptyset)&<Eu_2(\omega_2|\emptyset).
\end{aligned}$$

However, ω_1 is a perfect equilibrium, and thus a Nash equilibrium. Any behavioural strategy σ^i in the support of ω^i satisfies

$$Eu_i(\omega_1|\emptyset) \geq Eu_i((\sigma'', \omega^{-i})|\emptyset),$$

for any behavioural strategy σ''. Assume $\sigma'' = \omega_2$, then

$$Eu_i(\omega_1|\emptyset) \geq Eu_i(\omega_2|\emptyset).$$

It is contradicted to our assumption. Thus, a perfect equilibrium must satisfy the inequation (3).

Conclusions and Discussion

In this dissertation, we incorporate two different approaches to the analysis of finite decision problems and games with imperfect recall (including absentmindedness).

In the first chapter, we propose the psychological multiself approach. The approach transfers the decision problem from a one-stage procedure to a two-stage problem. The decision maker at the planning stage (ex-ante) chooses a strategy for the whole decision problem based on his beliefs about how he behaves at the execution stage (during execution). The decision maker is said to be confident if at the stage of planning, he believes he will follow whatever strategy he chooses now. The decision maker is said to be knowledgable if at the stage of planning, he knows he will reconsider the decision problem at each node and chooses a conditionally optimal one-shot behaviour. When he actually executes the decision problem, he chooses the conditionally optimal one shot behaviour at each of his reconsideration. Thus, the decision problem is described as a psychological game among multiselves of the decision maker. The results about equilibria indicate that a decision maker will never change his behaviour if he chooses an optimal plan before the decision problem starts. If the decision maker knows he might deviate from the current strategy he chooses, he would then actually deviate and behave differently from the situation if he believes himself.

The contribution of our approach is as follows. Firstly, we apply psychological games to analyse decision problems with absentmindedness. Psychological games normally model emotions or the interaction among individuals. It does not mean psychological games can only be used to model psychological factors. Strictly speaking, psychological games are used to model belief-dependent preference. The belief-dependent preference implies an individual's preference changes as his beliefs about others' behaviours change. In our approach, the planner's preference over the complete strategies of the decision problem also changes as his behaviour belief on the doers' behaviours. Thus, it is reasonable to incorporate psychological games into our approach. Furthermore, by different beliefs of the planner, we attempt to model the way of evaluating the decision problem ex-ante when a decision maker notices the appearance of future imperfect recall. In principle, it should be different from the way if the decision problem presents perfect recall. Besides those, our paper examines to what extent an ex-ante optimal strategy is conditionally optimal. The answer to it is that the ex-ante strategy is conditionally optimal even if the decision maker reconsiders the decision problem several times during one round of processing. The last point of contribution is that the strategy at an information set presenting absent-mindedness could determine the decision maker's position belief, and at the same time, the belief influences the decision maker's choice of strategy. The cycle of "strategy-belief-strategy" is unreasonable. However, in our approach, the doers are assumed to form position belief by the planner's strategy and it separates the process of position belief forming and the strategy choosing. By introducing a planner, the approach solves the cycle problem.

However, there are restrictions on our approach. Firstly, by introducing a planner into our model, we transfer the decision problem from a one-step procedure to a two-step procedure. Although our approach might be closer to description of the situations in real life, it increases the complexity of the decision making. In the aspect of computational algorithms, more data needs to be stored in our approach. We only propose two beliefs of the planner. There are many possibilities of beliefs to be explored. Despite those restrictions, our approach is a good attempt to analyse decision problems with absentmindedness.

In chapter 2 and 3, we develop the sophisticated recursive calculation rule and its resulting IS expected payoff function. The sophisticated recursive calculation rule describes the way how a sophisticated decision maker (player) considers the decision problem (game) at an information set. The decision maker (player) knows he will present imperfect recall at a future information set. Then, he takes how he assesses the decision problem at that information set into consideration when he assesses the decision problem at the current information set. Thus, for those outcomes which he does not know whether he can achieve at that information set, the current decision maker (player) still consider them even if he clearly knows some of them can never be reached. By recursion, the calculation rule states how the decision maker (player) evaluates his different behaviours which cause different outcomes.

In fact, the IS expected payoff at the information set containing the first node where the decision maker (player) acts is the same as the ex-ante expected payoff function in conventional decision (game) theory. It indicates that if we use IS expected payoff function as the standard to judge whether a strategy is time consistent (sequentially rational), then a strategy is ex-ante optimal if it is IS-time consistent (IS-sequentially rational). It indicates, that if the decision maker (player) is implementing a strategy which is conditionally optimal at any information set, then it must be the optimal strategy if he chooses ex-ante. In other words, if an individual does not have motivation to change his plan during execution of a decision problem (game), then he must have chosen the best plan before the decision problem (game) starts.

The main contribution of chapter 2 and chapter 3 is that the sophisticated recursive calculation rule provides a model to differentiate the decision maker (player)'s evaluation of the decision problem (game) between the situation that he knows he is imperfect recall and the situation that he does not know it. In conventional decision (game) theory, the two situations are modelled in the same way. Different from the psychological multiself approach in chapter 1, the decision maker (player) is allowed to choose his immediate behaviour and the behaviours at the information sets which can be reached or be imagined reached due to imperfect recall.

However, it seems to be a little complicated to develop the sophisticated recursive calculation rule. Although the story behind the calculation rule is reasonable and natural, the functional form of the calculation rule looks scary and there are numbers of necessary definitions that need to be defined due to the complexity of imperfect recall and the approach.

We use the ex-ante expected payoff function as the standard of a third party to judge whether the strategy that a decision maker (player) is implementing is a good choice or not. The approaches in the chapter 1,2 and 3 illustrate that preparing for the future imperfect recall is a good way to prevent a decision maker (player) changes to a suboptimal strategy during execution.

Appendix

Appendix I:

Belief System at an Information Set with Absentmindedness

The way defined in conventional game theory to calculate the belief system at an information set is not reasonable anymore if the information set presents absentmindedness. A belief system is a collection of probability measures μ. Denote $\mu(h|X)$ the probability of reaching each node $h \in X$ conditional on reaching an information set X. We say the probability of **reaching** a node is p, if a player with probability p passes the node when the game ends. Denote $p(h|\emptyset;\sigma)$ is the probability of reaching the node h by implementing the strategy σ.

In a game without absentmindedness, reaching each node in a non-singleton information set are mutually exclusive events. The probability of reaching an information set X is the sum of probability of reaching each node $h \in X$,

$$p(X|\emptyset;\sigma) = \sum_{h \in X} p(h|\emptyset;\sigma).$$

Then, the probability of reaching each node $h \in X$ conditional on reaching an information set X is

$$\mu(h|X;\sigma) = \frac{p(h|\emptyset;\sigma)}{p(X|\emptyset;\sigma)}.$$

We say the probability of **being** at a node is p, if with probability p a player is located at the node when the game is interrupted in the middle of execution. The events of being at different nodes are mutually exclusive under any circumstances. The probability of being at a node h and the probability of reaching the node h are identical if the following several assumptions are satisfied.

Assumption 3.9.1 *The game is finite and the longest branch of the game has m non-terminal nodes.*

Assumption 3.9.2 *Every player moves at the same speed. A player spends time t_0 to move one-step.*

Then, the game is located at the first node of each branch if the game is stopped at time $t \in [0, t_0]$, the game is located at the second node of each branch if the game is stopped at time $t \in [t_0, 2t_0]$, and so on.

Assumption 3.9.3 *For any time t_1 and t_2, the games is stopped by an external device at time $t_i, i = 1, 2$ with identical probability.*

It is possible that the game has ended when the external device is going to stop the game. The probability of being at one of the ith ($i \in \{1, ..., m\}$) nodes along branches is $1/m$. The event of being at the node h occurs, if the following two conditions are satisfied.

1. The game reaches the node h, the probability of it is $p(h|\emptyset;\sigma)$;
2. The game is stopped at the node h, the probability of it is $1/m$.

Thus, the probability of being at the node h is $p(h|\emptyset;\sigma)/m$. The probability of being at the node $h \in X$ conditional on being at the information set X is

$$\frac{p(h|\emptyset;\sigma)/m}{\sum_{h' \in X} p(h'|\emptyset;\sigma)/m} = \frac{p(h|\emptyset;\sigma)}{\sum_{h' \in X} p(h'|\emptyset;\sigma)} = \frac{p(h|\emptyset;\sigma)}{p(X|\emptyset;\sigma)} = \mu(h|X;\sigma).$$

Therefore, the relative probability of being at a node could be calculated by the probability of reaching different nodes.

However, if absentmindedness presents, the events of reaching different nodes in the same information set are not mutually exclusive. The probability of reaching the information set X is not necessarily equal to the sum of the probability of reaching each node in X. We use absentminded driver paradox (see Figure 3.1) to explain the situation of absentmindedness. Assume the decision maker's strategy is to continue with probability p. Denote $p(e_i), i = 1,2$ the probability of reaching the node e_i, $p(X)$ the probability of reaching the information set X. Then,

$$p(X) = p(e_1) \neq p(e_1) + p(e_2),$$

since the event of reaching the node e_1 must have occurred if the event of reaching the node e_2 occurs.

Besides the mutual exclusion issue, we also find the assumptions are not reasonable. Firstly, it is unrealistic to assume players make decisions at the same speed especially if more than one player is involved in the game. Secondly, the above assumptions illustrate that the game must be stopped by an external device. The external device might stop the game after the game ends. However, every player knows the fact that the game has not been completed when he is stopped during the execution of the game. Therefore, we assume that the one (an external device or a player) who takes charge of stopping the game knows the fact that the game is executing, but does not know the speed of making decisions by each player. Additionally, even for one player, the speed of decision making by a player is not necessarily to be always the same in a game. A player in an information set presenting absentmindedness does not know either how much time it takes from the beginning of the game or what is his speed of decision-making, if it is the player who stops the game. The same as the situation of no absentmindedness, we assume for any time t_1 and t_2, as long that the game has not ended, the probability of stopping the game at t_1 and t_2 are identical.

Discuss back to the absentminded driver paradox. In this example, the fact that the decision problem is executing is equivalent to the fact that the decision maker is at the information set X. The event of being at information set X contains and only contains three mutually exclusive events.

Event 1. Being at node e_1 and will move to e_2, denote p_1 the probability of event 1.

Event 2. Being at node e_1 and will move to e_2, denote p_2 the probability of event 2.

Event 3. Being at node e_2 denote p_3 the probability of event 3.

The event of reaching e_2 contains event 2 and 3. Thus,

$$p_2 + p_3 = p(e_2).$$

The relative probability of event 2 and event 3 is determined by the relative speed of the decision maker decides at node e_1 and e_2. If we assume the internal consistency of the decision maker, he would spend identical time on reconsideration at e_1 and e_2. Under this assumption,

$$p_2 = p_3 = p(e_2)/2.$$

The probability of event 1 is

$$p_1 = p(e_1) - p(e_2)$$

. Then, the probability of being at node e_1 conditional on being at information set X is

$$p_1 + p_2 = p(e_1) - p(e_2) + p(e_2)/2 = 1 - p/2.$$

The probability of being at node e_2 conditional on being at information set X is

$$p_3 = p/2.$$

Even without the assumption of internal consistency, we could calculate the probability of being at each intersection by introducing a parameter to describe the relative speed of reconsideration at node e_1 and e_2. Assume that, conditional on reaching the node e_2, the decision maker spends α portion of time to reconsider at e_1 and $1-\alpha$ portion of time to reconsider at e_2. Then, the probability of being at node e_1 conditional on being at information set X is

$$p_1 + p_2 = p(e_1) - p(e_2) + p(e_2)/2 = 1 - (1-\alpha)p.$$

The probability of being at node e_2 conditional on being at information set X is

$$p_3 = (1-\alpha)p.$$

We can see that, although the method to calculate the probability of being at each node cannot avoid the problematic issue of strategy-belief-strategy cycle, it solves the problem of mutual exclusion and relaxes the unrealistic assumptions of the game. The strategy-belief-strategy cycle indicates the decision maker's strategy determines his belief system on the information set with absentmindedness, and reversely, the decision maker's belief system influences strategy choosing. This problematic issue is solved in the psychological multiself approach by introducing the planning stage into the procedure of decision making.

Appendix II:

Figures and Tables

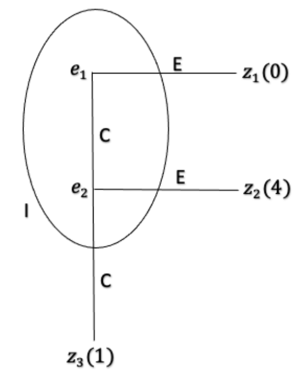

Figure 3.1: The Paradox of Absentminded Driver

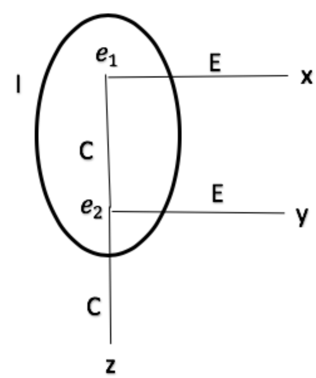

Figure 3.2: $x - y - z$ General Version

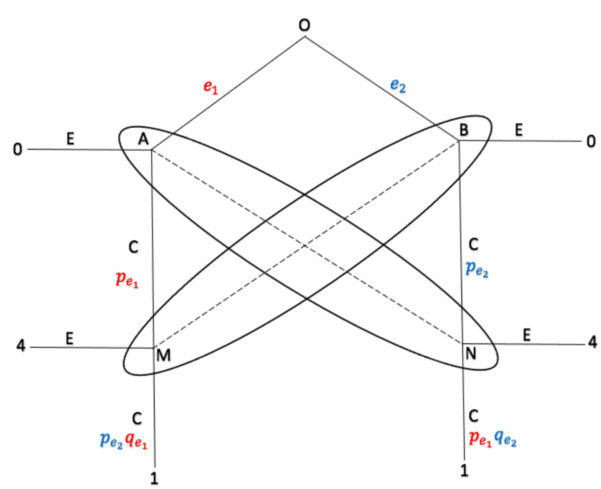

Figure 3.3: Identical and Symmetric

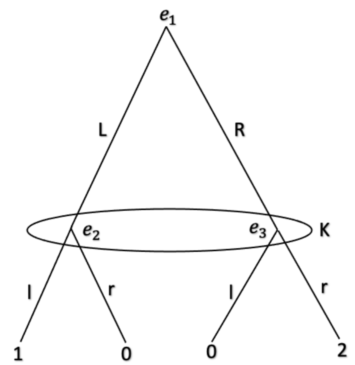

Figure 3.4: Forget Previous Moves

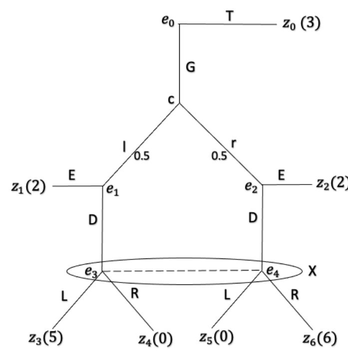

Figure 3.5: Forget Moves of the Chance Player

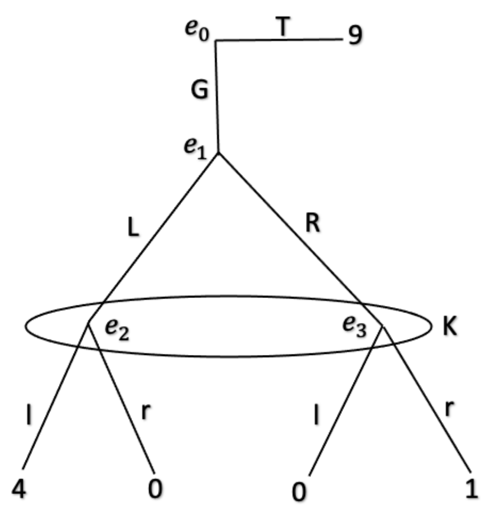

Figure 3.6

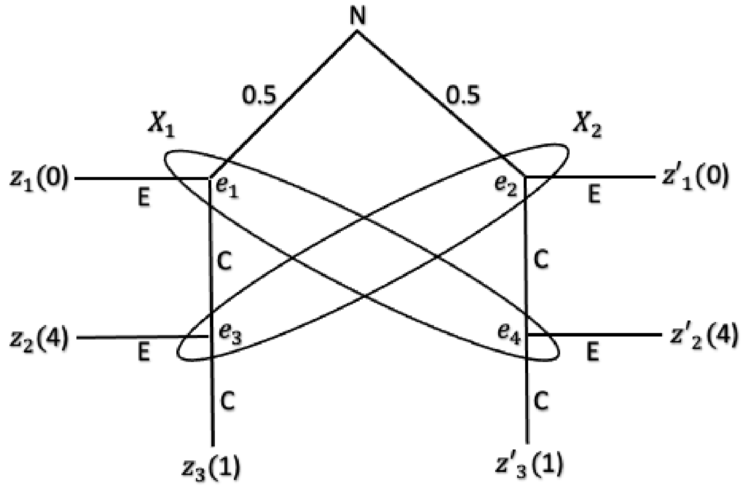

Figure 3.7: Gilboa Formation of Absentminded Driver Paradox

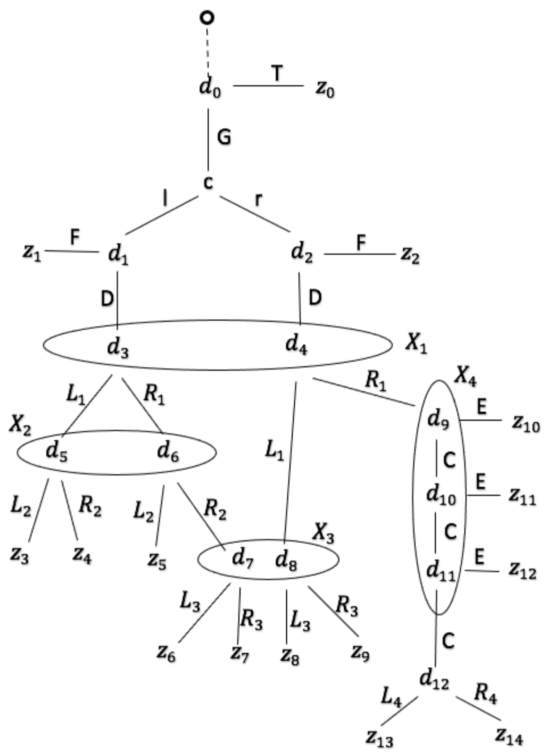

Figure 3.8: An Abstract Decision Problem

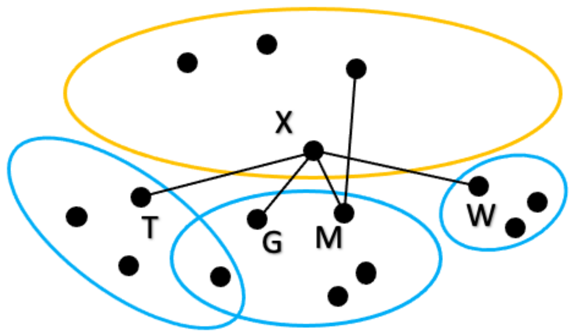

Figure 3.9: IS Time Consistency

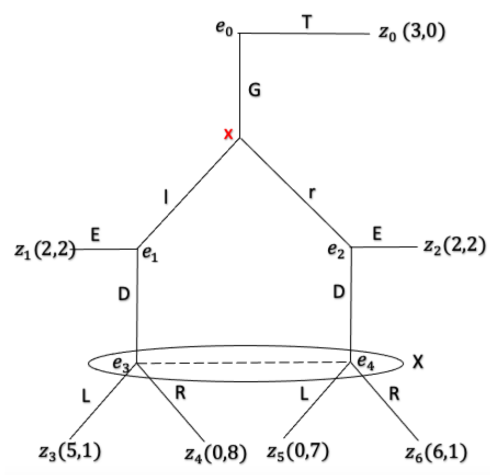

Figure 3.10: Forget Previously Acquired Information

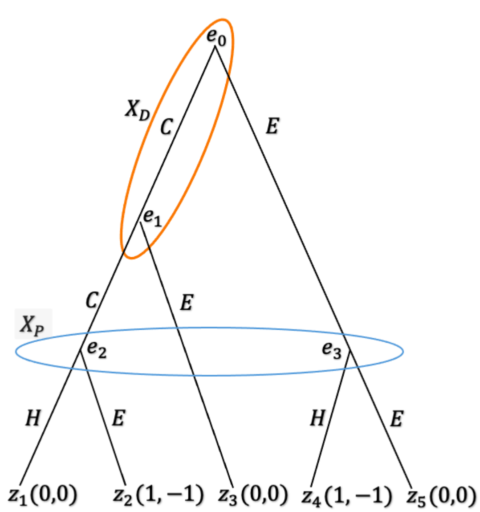

Figure 3.11: Absentmindedness

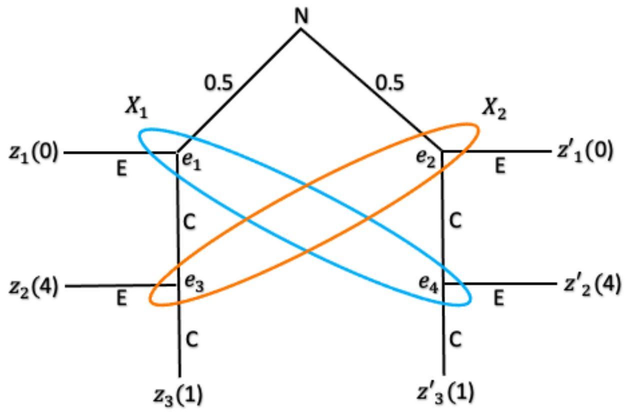

Figure 3.12: Cross-branch information sets

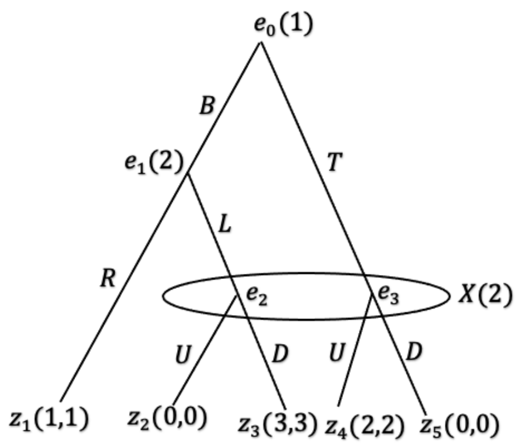

Figure 3.13: Comparison of sequential equilibria in terms of two expected payoff functions

	T	B
LU	(2,2)	(0,0)
LD	(0,0)	(3,3)
RU	(2,2)	(1,1)
RD	(0,0)	(1,1)

Table 3.1: Payoff table of player 1 and 2

??

Acknowledgement

I really appreciate the days I live in Glasgow and pursuing the degree, although the four-year Ph.D. life might be my hardest time so far. I'm writing this short statement to thank all the guys that I encountered during this period. Thank you to let me know that pursuing knowledge is full of touching, achievement, excitement, and at last, happiness.

First of all, I would like to give my sincere appreciation to my supervisors Prof. Sayantan Ghosal and Prof. Yiannis Vailakis. They are very nice advisors with an open mind, professional expertise, and kindly heart. They always give me support no matter in the scope of knowledge or inspirit, especially when I feel extremely frustrated.

Secondly, I would like to thank Dr. Theodore Koutmeridis. He was the course coordinator of Behavioural Economics. I took the tutorial work for this course for three years. We were such a good team to work together.

Then, I want to thank all the other colleagues and friends I meet in Glasgow. I really appreciate the time we spent together. I really appreciate Xiaoting Qin's mental support during the past six years. We have experienced much and she makes me feel I am not lonely.

Last but not least, I really appreciate the support from my parents and my grandmother. I might not make it without your encouragement and love.

Anyway, there are too many people to thank and acknowledge, too many things to appreciate, and too many staffs to remember. I would like to dedicate this thesis to all lovely lives. The past has gone, let's focus on the future. Start sailing, conquer the sea!

Best wishes to you all!

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Ruochen Dong

Bibliography

Alpern, S. (1988), 'Game with repeated decisions', *SIAM Journal on Control and Optimization* **26**(2), 468–477.
URL: *https://doi.org/10.1137/0326028*

Ambrus-Lakatos, L. (1999), An essay on decision theory with imperfect recall, KTK/IE Discussion Papers 1999/5, Budapest.
URL: *http://hdl.handle.net/10419/108019*

Attanasi, G. & Nagel, R. (2008), 'A survey of psychological games: Theoretical findings and experimental evidence', *Games, rationality and behavior. Essays on behavioral game theory and experiments* pp. 204–232.

Aumann, R. J., Hart, S. & Perry, M. (1997a), 'The absent-minded driver', *Games and Economic Behavior* **20**(1), 102–116.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905777*

Aumann, R. J., Hart, S. & Perry, M. (1997b), 'The forgetful passenger', *Games and Economic Behavior* **20**(1), 117–120.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905789*

Battigalli, P. (1997), 'Dynamic consistency and imperfect recall', *Games and Economic Behavior* **20**(1), 31–50.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905352*

Battigalli, P. & Dufwenberg, M. (2007), 'Guilt in games', *American Economic Review* **97**(2), 170–176.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.97.2.170*

Battigalli, P. & Dufwenberg, M. (2009), 'Dynamic psychological games', *Journal of Economic Theory* **144**(1), 1–35.
URL: *https://www.sciencedirect.com/science/article/pii/S0022053108000379*

Battigalli, P. & Dufwenberg, M. (2020), 'Belief-dependent motivations and psychological game theory'.

Battigalli, P., Dufwenberg, M. & Smith, A. (2019), 'Frustration, aggression, and anger in leader-follower games', *Games and Economic Behavior* **117**, 15–39.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825619300788*

Dimitri, N. (1999), 'Some comments on time-inconsistency in a game with absentmindedness', *University of Siena, Dept. of Economics Working Paper* (273).

Dimitri, N. (2009), 'Dynamic consistency in extensive form decision problems', *Theory and decision* **66**(4), 345–354.

Dufwenberg, M. & Gneezy, U. (2000), 'Measuring beliefs in an experimental lost wallet game', *Games and Economic Behavior* **30**(2), 163–182.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825699907157*

Dufwenberg, M. & Kirchsteiger, G. (2004), 'A theory of sequential reciprocity', *Games and Economic Behavior* **47**(2), 268–298.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825603001908*

Ederer, F. & Stremitzer, A. (2017), 'Promises and expectations', *Games and Economic Behavior* **106**, 161–178.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825617301677*

Falk, A. & Fischbacher, U. (2006), 'A theory of reciprocity', *Games and Economic Behavior* **54**(2), 293–315.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825605000254*

Fehr, E. & Schmidt, K. M. (1999), 'A theory of fairness, competition, and cooperation', *The Quarterly Journal of Economics* **114**(3), 817–868.
URL: *http://www.jstor.org/stable/2586885*

Geanakoplos, J. (1989), 'Game theory without partitions, and applications to speculation and consensus'.

Geanakoplos, J., Pearce, D. & Stacchetti, E. (1989), 'Psychological games and sequential rationality', *Games and Economic Behavior* **1**(1), 60–79.
URL: *https://www.sciencedirect.com/science/article/pii/0899825689900055*

Gilboa, I. (1997), 'A comment on the absent-minded driver paradox', *Games and Economic Behavior* **20**(1), 25–30.
URL: *https://EconPapers.repec.org/RePEc:eee:gamebe:v:20:y:1997:i:1:p:25-30*

Grove, A. J. & Halpern, J. Y. (1997), 'On the expected value of games with absentmindedness', *Games and Economic Behavior* **20**(1), 51–65.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905583*

Halpern, J. Y. (1997), 'On ambiguities in the interpretation of game trees', *Games and Economic Behavior* **20**(1), 66–96.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905571*

Halpern, J. Y. & Pass, R. (2016), Sequential equilibrium in games of imperfect recall, *in* 'Fifteenth International Conference on the Principles of Knowledge Representation and Reasoning'.

Hillas, J. & Kvasov, D. (2020a), 'Backward induction in games without perfect recall', *Games and Economic Behavior* **124**, 207–218.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825620301263*

Hillas, J., Kvasov, D. et al. (2020b), 'Kuhn equivalence of strategies'.

Isbell, J. (1957), 'Finitary games', *Contributions to the Theory of Games* **3**, 79–96.

Isbell, J. R. (1959), 'Absolute games', *Contributions to the Theory of Games* **4**, 357–396.

Kawagoe, T. & Narita, Y. (2014), 'Guilt aversion revisited: An experimental test of a new model', *Journal of Economic Behavior & Organization* **102**, 1–9.
URL: *https://www.sciencedirect.com/science/article/pii/S0167268114000596*

Khalmetski, K., Ockenfels, A. & Werner, P. (2015), 'Surprising gifts: Theory and laboratory evidence', *Journal of Economic Theory* **159**, 163–208.
URL: *https://www.sciencedirect.com/science/article/pii/S0022053115000873*

Kline, J. J. (2005), 'Imperfect recall and the relationships between solution concepts in extensive games', *Economic Theory* **25**(3), 703–710.

Kuhn, H. W. (1950), 'Extensive games*', *Proceedings of the National Academy of Sciences* **36**(10), 570–576.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.36.10.570*

Kuhn, H. W. (1953), *11. Extensive Games and the Problem of Information*, Princeton University Press, pp. 193–216.
URL: *https://doi.org/10.1515/9781400881970-012*

Kuhn, H. W. (2009), *Lectures on the Theory of Games (AM-37)*, Princeton University Press.
URL: *https://doi.org/10.1515/9781400829569*

Lambert, N. S., Marple, A. & Shoham, Y. (2019), 'On equilibria in games with imperfect recall', *Games and Economic Behavior* **113**, 164–185.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825618301490*

Lipman, B. L. (1997), 'More absentmindedness', *Games and Economic Behavior* **20**(1), 97–101.

Mertens, J.-F., Sorin, S. & Zamir, S. (2015), *Repeated games*, Vol. 55, Cambridge University Press.

Morgenstern, O. & Von Neumann, J. (1953), *Theory of games and economic behavior*, Princeton university press.

Morrison, W. & Rutström, E. E. (2002), 'The role of beliefs in an investment game experiment', *Available at SSRN 3140040* .

Osborne, M. J. & Rubinstein, A. (1994), *A course in game theory*, MIT press.

Piccione, M. & Rubinstein, A. (1997a), 'On the interpretation of decision problems with imperfect recall', *Games and Economic Behavior* **20**(1), 3–24.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825697905364*

Piccione, M. & Rubinstein, A. (1997b), 'The absent-minded driver's paradox: Synthesis and responses', *Games and Economic Behavior* **20**(1), 121–130.

Poddar, A., M Ellis, C. & Ozcan, T. (2015), 'Imperfect recall: The impact of composite spending information disclosure on credit card spending', *Journal of Consumer Policy* **38**(1), 93–104.

Rabin, M. (1993), 'Incorporating fairness into game theory and economics', *The American Economic Review* **83**(5), 1281–1302.
URL: *http://www.jstor.org/stable/2117561*

Rabinowicz, W. (2003), 'Remarks on the absentminded driver', *Studia Logica* **73**(2), 241–256.

Rubinstein, A. (1991), ‘Comments on the interpretation of game theory’, *Econometrica* **59**(4), 909–924.
URL: *http://www.jstor.org/stable/2938166*

Segal, U. (2000), ‘Don’ t fool yourself to believe you won?t fool yourself again’, *Economics Letters* **67**(1), 1–3.

Selten, R. (2020), 18. reexamination of the perfectness concept for equilibrium points in extensive games. int. j. game th. 4 (1975) 25-55., *in* ‘Classics in Game Theory’, Princeton University Press, pp. 317–354.

Strotz, R. H. (1955), ‘Myopia and inconsistency in dynamic utility maximization’, *The Review of Economic Studies* **23**(3), 165–180.
URL: *http://www.jstor.org/stable/2295722*

Wichardt, P. C. (2008), ‘Existence of nash equilibria in finite extensive form games with imperfect recall: A counterexample’, *Games and Economic Behavior* **63**(1), 366–369.
URL: *https://www.sciencedirect.com/science/article/pii/S0899825607001534*