UNIVERSITY OF NOTTINGHAM

DISCUSSION PAPERS IN ECONOMICS

No. 98/11

Extensive Games of Imperfect Recall and Mind Perfection

Matt Ayres

Abstract

In this paper we examine how the addition of imperfect recall as a perturbation to a perfect recall game can be used as an equilibrium refinement. We discuss the properties of two such concepts, from the addition of complete confusion between similar histories to considering small trembles'in a players beliefs. Central to our discussion is the notion of which decisions can reasonably be confused and we suggest that modelling informational confusion may be a useful way of measuring strategic complexity.

1. Introduction

Game theoretic modelling of economic agents as rational players has led to paradoxes and serious discrepancies between observation and theory. One way we might hope to gain further insights into behaviour and existing results is by adding psychological elements to the reasoning process of players. Such models are in the class of those dealing with ‘bounded rationality’.[1] The simple psychological addition we consider here is to model players with ‘bad memories’ or imperfect recall.

Due to recent work by Piccione and Rubinstein (1997), games and decision problems with imperfect recall have been re-examined. This re-examination has indicated difficulties both in modelling and the possibility of a new type of time inconsistency problem, as demonstrated by their example of the ‘absentminded driver’. Most of the papers following Piccione and Rubinstein, have concentrated on this time consistency problem, suggesting a variety of interpretations and resolutions (Aumann et al (1997a, 1997b), Battigalli (1997), Binmore (forthcoming), Gilboa (1997), Grove and Halpern (1997), Halpern (1997) and Lipman (1997)).

In this paper we take a different approach, introducing imperfect recall as a perturbation to a perfect recall game and examine how the addition of ‘bad memory’ affect the set of equilibrium predictions. Our rationale is that players should, ceteris paribus, prefer equilibria in which they use ‘less cognitively demanding’ strategies, in our case ones that require ‘less memory’. Note the link between memory limitations and strategic complexity has been made in the literature on finite automata play in infinitely repeated games, (for example, Abreu and Rubinstein (1988) and Rubinstein (1987)).

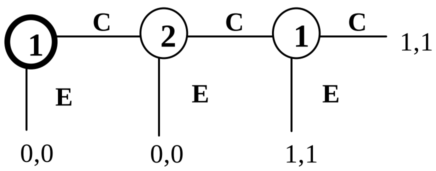
Our first refinement, which we call ‘mind perfection’, deals with a complete reduction of a player's ability to distinguish between similar decision nodes. By introducing as much imperfect recall and imperfect information as possible we ask the question ‘Which, if any, equilibria survive?’ Thus, a mind perfect strategy corresponds to a very simple strategy in which the player does not have to distinguish similar decision nodes.

---

[1] The general acceptance of bounded rationality in economics is largely is due to the work of Simon (1978 and others). In particular Simon’s emphasises the potential linkages between economic and psychological research.

As an example, consider the perfect recall game in figure 1[2]. A rational player 1 would continue at the first decision node and at the third decision node he is indifferent between actions *E* and *C* (both *{C,C,E}* and *{C,C,C}* are subgame perfect). Under mind perfection, we assume that player 1 confuses the similar decisions he makes at the first and third decision node. A decision node or history is said to be similar to another if it satisfies the minimal requirements for being in the same information set.[3] For an absentminded player 1, confusing the first and third decision node, action *C* is strictly better than action *E*. Our interpretation is that for player 1 the strategy of always continuing *{C,C}* requires less memory than the strategy *{C,E}*, i.e. strategy *{C,C}* is less 'complex' than strategy *{C,E}*. Fundamental to this example and the rest of this paper is the notion of which decision nodes (histories) can reasonably be confused, a notion we describe as 'similarity'.

Figure 1



Mind perfection is a very strict requirement and consequently it lacks application in all but a few games. However, a mind perfect equilibrium is stable even when the game is perturbed to the extent that all similar histories are in the same information set. Following Selten's (1978, p147-152) distinction, we wish to classify solutions that require only minimal information as corresponding to problem of a 'routine' nature. An equilibrium which fails to meet the requirements of a mind perfect equilibrium needs application of some further 'reasoning'.

---

[2] The initial node, where *h*=˘ , is represented by a bold circle.

[3] As a minimum we require that players always know the set of actions from which they are choosing and whose move it is.

Our second refinement considers the possibility of imperfect recall and imperfect information as a small perturbation in a player's belief about what decision is being made. Equilibria which survive such perturbations we call 'trembling mind perfect', since in a similar way to trembling hand (perfect) equilibrium we wish to find equilibria which survive the addition of 'mistakes'. The rationale here is that these mistakes are based around perturbations in beliefs (mistakes confusing similar histories) rather than mistakes in taking actions. Intuitively we find mistakes based on perturbations in beliefs to be more appealing since they are determined by the player's perception of the games structure, i.e. we can provide a psychological explanation.[4]

Section 2 introduces the required notation and formal definitions of imperfect recall. Section 3 defines our notion of similarity and notes the problems in such an explicit definition. Our two equilibrium refinements and discussion of their properties are given in section 4. We also note the concept of a 'trembling mind' and
independent.

2. An Extensive Game with Imperfect Recall

We define a finite extensive game Γ with imperfect information and imperfect recall as

$$\Gamma = < H, N, P, \succsim_i, r(h), A(h)_{h\in H}, (I_i)_{i\in N}, X_i(h)_{h\in H, i\in N} >$$

where:

H is a set of finite histories, such that $\emptyset \in H$ and if a sequence of actions $(a^k)_{k=1....K} \in H$ then $(a^k)_{k=1....L} \in H$ for all $L<K$. All histories begin with $\emptyset$ and contain the moves made in sequential order. A history is interpreted as a physical description of all the moves made by the players (including chance) and we label each node with this history. We define Z to be the set of terminal histories, where a history is said to be terminal if there is no a^{K+1} such that $(a^k)_{k=1....K+1} \in H$. The node with a history of just $\emptyset$, is called the initial node. Graphically we represent the initial node of any game with a bold circle.

---

[4] Myerson's (1978) Proper equilibrium takes a different approach to rationalising trembles, but mistakes still remain based on actions rather than beliefs.

N is a finite set of players, not including the chance (nature) player which we denote *c*. *P* the player function assigns a player in $N \cup \{c\}$ to move after every non-terminal history *H\Z.* For each player *i* in *N* there is a preference relation $\succeq_i$ on *Z.*

When *P(h)=c* the next move is made by the chance player and $\rho(h)$ assigns a probability of occurrence to each action $a \in A(h)$.

After every non-terminal history $h \in H \setminus Z$ player *P(h)* chooses an action from $A(h)=\{a:(h,a) \in H\}$. To avoid degeneracy we assume *A(h)* contains at least two elements.

For each player $i \in N$ there is a partition $\boldsymbol{I_i}$ of histories $h \in H$ at which *P(h)=i*. For each $I_i \in \boldsymbol{I_i}$ (an information set), any two histories h and h' in I_i must satisfy the property that $A(h) = A(h')$. We label the actions available at the information set I_i, as $\overline{A}(I_i)$ such that $\overline{A}(I_i) = A(h)$ for all $h \in I_i$. The interpretation of an information set is that all histories (nodes) in I_i cannot be distinguished from one another. Graphically we depict information sets as a shaded box linking two or more decision nodes.

$X_i(h)$ is the players experience at the history *h*, consisting of a pair $\{X_i^I(h), X_i^a(h)\}$ such that:

$X_i^I(h)$ is a sequential ordering of information sets player *i* has visited in order to reach *h* (player *i*'s experience of information sets at the history *h*).

$X_i^a(h)$ is a sequential ordering of actions player *i* has taken in order to reach *h* (player *i*'s experience of actions at the history *h*).

Using our definition of an extensive game we can define situations of imperfect recall.

Definition: Imperfect Recall

A player i is said to have imperfect recall if $X_i(h) \neq X_i(h')$ *for any histories h and* h' *that are in the same information set* I_i.

In line with the Piccione and Rubinstein (1997), we distinguish three types of recall problems. **Imperfect Recall of Information Sets:** where a player forgets the sequence of information sets through which play has past. **Imperfect Recall of Actions:** where the player recalls he has made a prior move but not what action he chose. **Absentmindedness:** where the player cannot recall whether he has made a prior move or not.

More formally,

Definition: Imperfect Recall of Information Sets

If for some h and h' *in the same information set,* $X_i^I(h) \neq X_i^I(h')$ *then player i is said to have imperfect recall of information sets.*

Definition: Imperfect Recall of Actions

If for some h and h' *in the same information set,* $X_i^a(h) \neq X_i^a(h')$ *then player i has imperfect recall of actions.*

Definition: Absentmindedness

If for some h and h' *in the same information set,* $a \in A(h)$ *is part of the sequence of actions* $X_i^a(h')$ *then player* i *is said to suffer from absentmindedness.*

For h and h' in the same information set, the statement $a \in A(h)$ is part of the sequence of actions $X_i^a(h')$ is equivalent to *h* being a subhistory of *h'*. Thus the definition presented above is equivalent to that of Piccione and Rubinstein.

Definition: Absentmindedness (Piccione and Rubinstein)

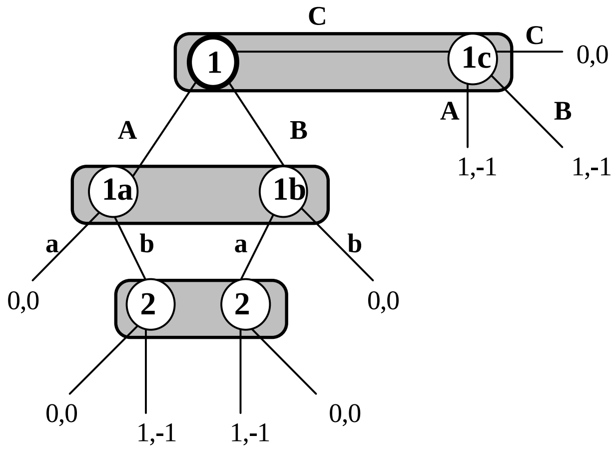
A player i is absentminded if for some h and h' *are in the same information set where* $h'=(a^k)_{k=1....K}$¿ *H and* $h=(a^k)_{k=1....L}$¿ *H for L<K.*

Absentmindedness represents a special case of imperfect recall where the player fails to recall both visiting an information set and the action he took there. From our definition of absentmindedness since $a \in A(h)$ is part of the sequence of actions $X_i^a(h')$ the sequence of information sets must be such that $X_i^I(h) \subset X_i^I(h')$ and hence $X_i^I(h) \neq X_i^I(h')$ (our definition of imperfect recall of information sets).

Since, under absentmindedness we allow h' to be a subhistory of *h*, for h and $h \in {}_i$, the same information set can be visited more than once. Whilst we allow an information set to be visited more than once it must be to a different decision node, i.e. we do not allow infinite cycles or H (the set of histories) to be infinite.

Figure 2 shows a game with both absentmindedness and imperfect recall of actions. At his first decision node player 1 suffers from absentmindedness, that is he is unsure whether he is indeed at the initial node or he has already chosen the action *C* and is at node *1c*. The player is similarly confused at node *1c*. In the information set joining nodes *1a* and *1b* player 1 has imperfect recall of actions, that is he knows he has chosen either *A* or *B* at the initial node but cannot recall which.

Figure 2: An Extensive Game with Absentmindedness and Imperfect Recall of Actions (Isbell 1957, p85)



It should that whilst imperfect recall is clearly defined in the extensive form, many definitions preclude cases of imperfect recall due to the difficulties it imposes. For example, Kreps and Wilson's (1982) sequential equilibrium extends only to cover perfect recall games.

Choice of strategy in games of imperfect recall is problematic because the possible structure of information sets renders some results concerning strategic equivalence invalid.

Definition:

*A **pure strategy** for player $i \in N$ in an extensive game is a function assigning a single action in $\overline{A}(I_i)$ to each information set $I_i \in$ $\boldsymbol{I_i}$. A **mixed strategy** for player $i \in N$ is a probability measure p_i over the set of player i's pure strategies. A **behavioural strategy** for player $i \in N$ is a collection of independent probability measures $b_i(I_i)$ for all $I_i \in$ $\boldsymbol{I_i}$, where $b_i(I_i)$ assigns a probability to each action in $\overline{A}(I_i)$.*

We interpret a pure strategy as a plan of action formulated before play begins. A mixed strategy is thus a randomisation over such plans of action and a behavioural strategy is a single plan of action, with instructions for the player to randomise on reaching an information set. In the case of absentmindedness we allow the player to visit an information set more than once in any play of the game. The player is restricted to following the same behavioural randomisation at all decision nodes in the information set and this randomisation is assumed to be realised each time he/she visits that information set. Further discussion of this point can be found in Piccione and Rubinstein (1997).

In games involving imperfect recall behavioural and mixed strategies are not outcome equivalent (see for example, Osborne and Rubinstein 1994, p.203-204). In this paper we restrict strategy choice to behavioural strategies. We ignore mixed strategies without behavioural equivalents since they necessarily involve the player recalling the result of a centralised randomisation chosen at the beginning of the game. Such a device can be used to overcome some of the problems of imperfect recall. Since behavioural strategies are defined at the level of the information set, no such problem arises.

3. Similarity

Our formal definition of (strong) similarity, states that one history is similar to another if they satisfy the minimal requirements for being in the same information set. We are only interested in the number of distinct similar histories and thus define similarity to exclude a history being similar to itself. Further we assume that no history at which nature moves can be similar to any other. The latter requirement is imposed since we find no clear interpretation of what it means for nature to become 'confused'.

Definition: (Strongly) Similar Histories

A history $h \in H \setminus Z$ is similar to the history $h' \in H \setminus Z$ if and only if

$h \neq h'$ *and*

$A(h) = A(h')$ *and*

$P(h) = P(h') \neq c$.

In this paper we only consider similarity as defined above, but note that deciding what information is relevant for decisions to be classed 'similar' is not an easy task.. Firstly, we require that actions are not solely defined in terms of what choices are taken (e.g. whether to go left or right) but also reflect the state in which the choice is made (e.g. whether it is light or dark). Thus if a player is able to distinguish between light and dark, the action 'left when light' and 'left when dark' should be labelled differently in the extensive form (i.e. the decisions cannot be similar).

Secondly, if we allow players to have small doubts about their memories rather than complete memory loss, it may be more appropriate to consider a wider range of possible confusions. For example, consider that if at one decision node a boundedly rational player chooses between a set of actions *{a,b,c}* and at another *{a,b,c,d}*. We may wish to model the situation where on arriving at the first decision node the player has some positive belief he is at the second decision node but has failed to notice the action {d}.

4. Equilibrium Refinements

For simplicity we state our formal definition of mind perfection as a refinement of subgame perfection (i.e. applicable to perfect information games). Let *S(h)* be the set of histories strongly similar to *h*.

Definition: Subgame Perfect Equilibrium (SPE)

For Γ *a game with perfect recall and perfect information, let* $\Gamma(h)$ *be a subgame of* Γ *beginning at the history h. A subgame perfect equilibrium is a strategy profile* s^* *(consisting of a pure strategy* s_i^* *for each player i) which, for all histories* $h \in H \setminus Z$*, the strategy profile beginning at h, is a Nash equilibrium of the subgame* $\Gamma(h)$*.*

Definition: Mind Perfect Equilibrium (MPE)

For Γ *a game with perfect recall and perfect information, a mind perfect equilibrium is one which is both subgame perfect and satisfies the property that if* h and h' *are similar histories then* $s_i^*(h) = s_i^*(h')$*, where* $s_i^*(h)$ *is the choice of action in the pure strategy* s_i^* *at the history h.*

We consider the following properties of mind perfection in game Γ with perfect recall and perfect information. As preliminary we define:

Definition: Strict Dominance in Similar Histories

An action $a \in A(h)$ *is said to be strictly dominant in histories similar to h if in at least one history* $h' \in S(h) \cup h$, $a \succ_i a'$ *and for all other histories* $h'' \in S(h) \cup h$, $a \succeq_i a'$*, for all* $a' \in A(h)$ and $a' \neq a$.

We obtain the following results.

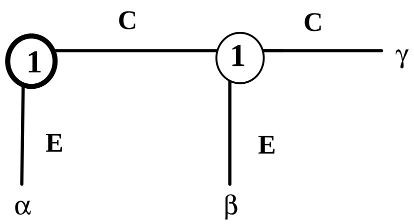
Theorem 4.1 All games Γ that contain no similar histories have at least one mind perfect equilibrium.

Proof: If Γ all histories $h \in H \setminus Z$, $S(h) = \varnothing$ by definition mind perfection and subgame perfection are identical. Since every game has at least one subgame perfect equilibrium, if there are no similar histories, every game must have a mind perfect equilibrium. ÿ

Theorem 4.2 In a game Γ with at least one pair of similar histories there may exist one mind perfect equilibrium, there may exist no mind perfect equilibrium or there may exist multiple mind perfect equilibria.

Proof: Consider the single player extensive form game (a decision problem) in figure 3. All three histories at which the player moves have similar histories. If the payoffs are such that $g > b$ and $g > a$ then there is a unique mind perfect equilibrium *{C,C}*. In the case where $b > g$ and $g > a$ then there are no mind perfect equilibrium. If $a = b = g$ then there are two mind perfect equilibria *{C,C}* and *{E,E}*. Hence there may exist a unique, there may exist no or there may exist multiple mind perfect equilibria. ÿ

Figure 3



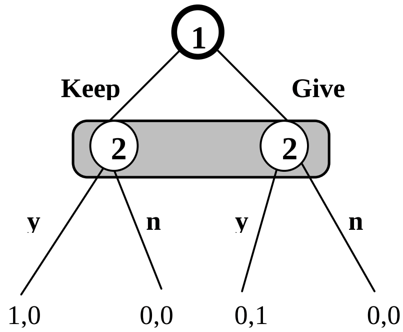
Theorem 4.3 A game Γ has a unique mind perfect equilibrium if for every history, (including those where $S(h) = \varnothing$), there exists an action which satisfies strict dominance in similar histories.

Proof: In the case where Γ a game with perfect recall and perfect information has no similar histories, $S(h) = \varnothing \ \forall h$. If strict dominance in similar histories is satisfied at the history *h*, there exists a unique $a \succ_i a'$ for all $a' \in A(h)$ and $a' \neq a$. Since at each history there is a single dominant action there is a unique subgame perfect equilibrium and by 4.1 a unique mind perfect equilibrium.

In the case where Γ, a game with perfect recall and perfect information, has at least some histories similar to $h \in H \setminus Z$. For each player i we assume for all histories h strict dominance in similar histories is satisfied by some action $a \in A(h)$. This action must be unique for all $h' \in S(h) \cup h$, since for some h', $a \in A(h') \succ_i a' \in A(h')$ for all $a' \neq a$ excludes $a' \in A(h) \succ_i a \in A(h)$ for any $a' \neq a$. The action $a \in A(h')$ must be part of a strategy s_i^* that for each player i makes up a subgame perfect equilibrium of Γ (since a is either strictly or weaker dominant at h'). Since this action is unique and the same at all histories similar to h, s_i^* will also be part of the unique mind perfect equilibrium of Γ. ÿ

As an example we consider the ultimatum game with a single indivisible good as shown in figure 4. In this game, player 1 can choose to give the good to player 2 or choose to keep it for himself, where the payoffs reflect the final holdings of the good. This game has two subgame perfect equilibria *{keep,y}*, *(keep,n}* and a single mind perfect equilibrium *{keep,y}*. Whilst subgame perfection removes the possibility of incredible threats (such as player 2 choosing *{n}* following *{give}*), mind perfection can be interpreted as removing weak credible threats. Mind perfection uses strict preferences in similar parts of the game tree to rule out some weakly dominated actions (a weakly credible threat). In the example in figure 4 player 2's dominant action of *{y}* following *{give}* and indifference between *{y}* and *{n}* following *{keep}*, result in the single mind perfect outcome *{keep,y}*. This strategy does not require player 2 to distinguish his two decision nodes. Consequently, such a strategy may be seen as one of lower complexity and which has a lower memory requirement (player 2 does not have to remember what player 1 chose).

Figure 4



Although we have stated mind perfect equilibrium as a refinement of subgame perfection, we can use the same idea to form a criterion that can be applied to any strategy (possibly behavioural) or equilibrium set of strategies.

Definition: Mind Perfection Criterion

A strategy b_i is said to be mind perfect if for all histories h and h' which are similar, $b_i(h) = b_i(h')$, where $b_i(h)$ is behavioural action in the strategy b_i at the history *h*.

Thus taking the example given in figure 3, the strategies *{E,E}*, *{C,C}* and behavioural randomisation assigning the same probability to taking action *{C}* at both decision nodes satisfies the mind perfect criterion. In the case where $a > b, g$ and $g > b$, the strategy *{E,E}* is not a mind perfect equilibrium according to our definition, it is, however, a Nash equilibrium which satisfies the mind perfect criterion.

Our second concept looks at the situation where each player assigns a positive, but possibly small probability e to confusing similar histories. Equilibrium strategies which survive such perturbations are said to be trembling mind perfect. As a preliminary we define a game involving such perturbations.

Definition: The perturbed game Γ_i'

For a finite extensive game Γ *with perfect recall, let* Γ_i' *be perturbation of* Γ *where for all histories* $h \in H \setminus Z$ *and P(h)=i, nature causes player* i *to believe they are at a history* h' *when he/she is at the history h, with probability* $\varepsilon_h^{h'}$ *, such that*

$$\varepsilon_h^{h'} > 0 \text{ for all } h' \in S(h),$$

$$\sum_{h' \in S(h)} \varepsilon_h^{h'} < 1$$

and, for simplicity, we assume that play can never reach a path in which player i makes more than one mistake.

Thus any path of the perturbed game where *h* is confused with a history h' and *h* is a subhistory of h' will be one involving a situation of absentmindedness. Any path where *h* is confused with a history h' such that *h* is not a subhistory of h' and h' is not a subhistory of *h* will be one involving either imperfect recall of actions.

Using our definition of a perturbed game we can define a trembling mind perfect equilibrium.

Definition: Trembling Mind Perfection

Let β_i^* *be a behavioral strategy for player i that is a best response given the actions of the other players* β_{-i}^* *in the unperturbed game* Γ*. The behavioral strategy* β_i^* *said to trembling mind perfect if there exists a perturbed game* Γ_i' *in which* β_i^* *is also a best response to* β_{-i}^**. A trembling mind perfect equilibrium consists of a behavioral strategy for each player i which is a best response in the game* Γ_i'*.*

Note that the perturbed game Γ_i' consists of one path where player *i* makes no mistakes, occurring with probability $\left(1 - \sum_{h' \in S(h)} \varepsilon_h^{h'}\right)$ and a further *n* paths for each history *h*, that has *n* similar histories *h'* each occurring with probability $\varepsilon_h^{h'}$ (i.e. *n* is the cardinality of the set *S(h)*). For example: Consider the one player extensive game (decision problem) Γ, shown in figure 5, with a single pair of similar histories *h* (the players first move) and h' (the player's second move). This gives a perturbed game Γ_i' (figure 6) where nature selects between three possible paths. One path where no mistakes are made (occurring with probability $(1-e_{h'}^{h}-e_{h}^{h'})$), one when reaching *h* the player thinks he is at h' (occurring with probability $e_{h}^{h'}$) and one when reaching h' the player thinks he is at *h* (occurring with probability $e_{h'}^{h}$). Note in the dark shaded information set the player believes he is at *h* and in the light information set thinks he is at h'.

A game Γ with two histories h' and h'' both similar to h and all other histories distinct (non-similar) gives a perturbed game where nature selects between seven possible paths.

Figure 5

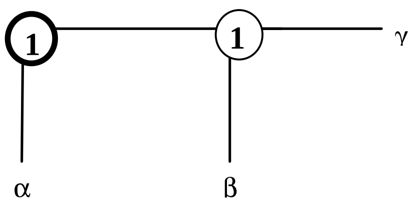
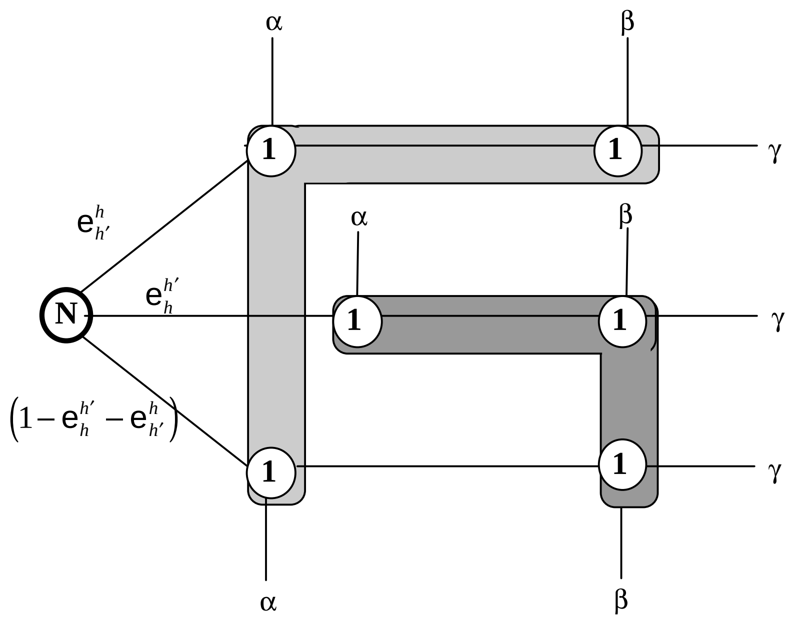


Figure 6



From our definition of trembling mind perfection, it follows:

Theorem 4.4 For any perfect recall game Γ there exists at least one trembling mind perfect equilibrium.

The proof of Theorem 4.4 is given in the Appendix

From Theorem 4.4 it follows that any extensive game with perfect recall has at least one trembling mind perfect equilibrium. Note that if Γ contains similar histories and there are two or more equilbria in behavioral strategies, equilbria will be excluded from being perfect to a trembling mind if they involve playing actions that are strictly dominated in similar histories. More formally,

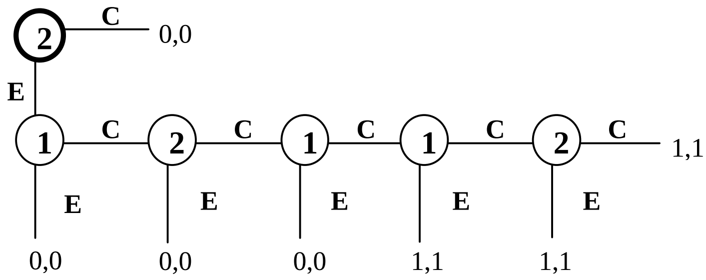
Theorem 4.5 Strategies which involve playing actions which are strictly dominated in similar histories will not be trembling mind perfect.

Proof: Consider that under an equilibrium strategy β_i^* the action *a* is chosen at the history h and that *a* is strictly dominated in similar histories by the action a'. In particular suppose, that $a' \succ_i a$ at history h', where $h' \in S(h)$. Then it follows that *a*, cannot be part of a trembling mind perfect strategy since in all possible perturbations confusing *h* and h' with positive probability following a' is strictly better. ÿ

It should be obvious that all mind perfect strategies will also be trembling mind perfect. Consider two examples. The extensive game shown in figure 1 the equilibrium in which player 1 plays *{C,C}* and player 2 plays *{C}* is both the unique mind perfect and unique trembling mind perfect equilibrium.

The extensive game shown in figure 7 is more interesting. Players making repeated choices between *C* and *E,* all histories for player 1 are similar to three other histories and all histories for player 2 similar to two other histories. Note that player 1ś second and third move immediately follow one another. Equivalence principles suggest that such moves could be coalesced into a single choice i.e. player 1 selects between the actions *E*, *CE* and *CC*. It should be obvious that in this paper we require such decisions to remain separate so as to allow the possibility of confusion between these two moves.

Figure 7



The game in figure 7 has no mind perfect equilibria (since no action strictly dominates in similar histories for player 2) and two trembling mind perfect equilibria *{E,C,C,C,C,C}* and *{E,C,C,C,C,E}*. Note that for some values of $e_h^{h'}$ player 2 prefers to play C and for other values prefers E, thus both equilibria are trembling mind perfect. Note also that player 1 strictly prefers to play C at his last decision node given he holds some positive belief about being at a similar history.

Finally it should be noted that not all trembling mind perfect equilbria are trembling hand perfect and not all trembling hand perfect equilbria are trembling mind perfect. Note that in the case where a game has a unique Nash equilibrium, it must also be trembling hand perfect, sequential and trembling mind perfect.

The example in figure 8 has a unique trembling mind perfect equilibrium *{R,r}* and two sequential equilibria, *{L,l}* and *{R,r}*. Whilst *{L,l}* is sequential it is not a trembling hand perfect equilibrium. This is one of the special cases where trembling hand and sequential equilibrium diverge. Both concepts have a consistency requirement, but only the former considers trembles' (see Kreps and Wilson, 1982 p882). Our concept of trembling mind perfection does not have this same consistency requirement but does introduce trembles.

If we change the game depicted in figure 8 so that player 2's decision nodes are all singletons, *{L,l}* is still not a trembling mind perfect equilibrium although it is both a sequential equilibrium and is trembling hand perfect. Similarly in figure 1, *{C,C,C}* and *{C,C,E}* are trembling hand perfect but only *{C,C,C}* is trembling mind perfect. Not all trembling hand perfect equilibria are trembling mind perfect. This is because trembling mind perfection allows for errors to be made between decision nodes that are not in the same information set but satisfy our definition of similarity.

Figure 8

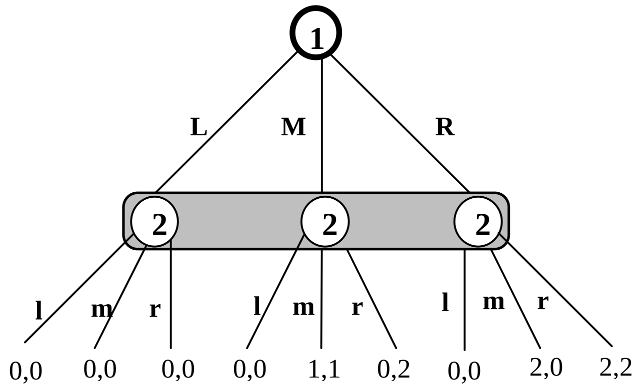
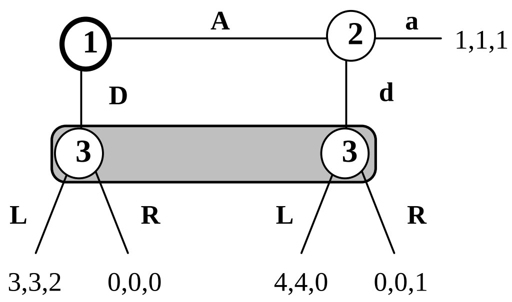


Figure 9 (Selten's Horse) shows a three player game. Play begins at the top left node and continues either across to player 2's node or downwards to player 3's information set. This example illustrates that not all trembling mind perfect equilibrium are trembling hand perfect. The unique trembling hand perfect (and sequential) equilibrium of the game is *{A,a,r}*, where *r* is played with at least a probability of three quarters. However, *{D,a,L}* is also a trembling mind perfect equilibrium. Note that, we do not claim that this last equilibrium is necessarily sensible, merely that it cannot be ruled out by a trembling mind (although it can by a trembling hand).[5]

[5] Since {D,a,L} is a Nash equilibrium which does not involve player 3 choosing an action which is strictly dominated in similar histories and hence it must be trembling mind perfect. Note that players 1 and 2 have no similar histories so for them this requirement is trivially satisfied.

Figure 9



We summarize the relationships between equilbria in the following theorem,

Theorem 4.6 There exist extensive games such that**:**

I. not all trembling mind perfect equilibria are trembling hand perfect
II. not all trembling hand perfect equilibria are trembling mind perfect
III. those equilibria which are trembling hand perfect but not sequential equilbria, are trembling mind perfect.

5. Concluding Remarks

We find the concepts presented in this paper to be appealing in that they use stability to the addition of confusion (mistakes in beliefs) as a selection criterion. Part of this appeal is that our models can be interpreted as giving a 'psychological' explanation of mistakes. The extreme case of mind perfection corresponds to limiting the players to very simple strategies. The problem with such severe restriction on strategy choice is that is possibly further from a desired model of descriptive behavior than is that of the fully rational player. At the other extreme we have considered strategies that are perfect to a trembling mind. By adding only an arbitrarily small amount of confusion we can guarantee existence, although the power of the concept is limited to removing only some weakly dominated strategies.

Thus whilst we find the idea of confusion to have psychological appeal, we do not see either of our concepts as presenting a model that is sufficiently rich to be appealing in describing actual behaviour. The purpose of considering the extremes of mind and trembling mind perfection is largely instructive; we consider players should play mind perfect strategies where possible and avoid those which are not trembling mind perfect. To move towards a model with more descriptive appeal we need to consider levels of 'confusion' between our two extremes. In such a model the confusion or the fear of confusion between similar decisions becomes an integral part of the decision making process. We also suggest that such a model may give us a way of ranking strategies on grounds of complexity. A simple strategy being one which is optimal irrespective of the players fear about becoming confused in carrying out his strategy.

It is also important to note that the power of our concepts is dependent upon our definition of similarity, i.e. what histories can reasonably be confused. There is unlikely to be any single correct definition of similarity, but rather it is likely to depend upon the situation being modeled and in particular the player's perception of the different decisions being made.

Finally, all the models considered here have assumed memory limitations are exogenously determined. An alternative would be to try formally incorporate the costs of recall into the games structure. Such models would allow players a more active cognition and seem appealing if the costs of recall are easy to identify, for example, where players are firms we may imagine recall costs as the cost of searching through a filing system or database to extract historic information. Our primary concern in this paper has been with individual human decision making agents where problems of recall are real but the costs of such are not easy to quantify.

Appendix

To show the existence of trembling mind perfection we characterize three situations (i-iii). The first is trivial, formal proofs of the other two are given below.

(i). The game Γ has no similar histories, thus for all players *i* the games Γ and Γ_i' are identical. All equilibrium strategies b^* in Γ are trembling mind perfect.

(ii). The game Γ contains similar histories and a unique equilibrium in which each player *i* follows the behavioral strategy b_i^* which is a best response to the strategies of the other players b^*_{-i}. This will also be the unique best response for the game Γ_i' and hence trembling mind perfect for the game Γ.

Proof: We want to show for all players *i*, that there exists an $e_h^{h'} > 0$ (and $\sum_{h' \in S(h)} e_h^{h'} < 1$) such that b_i^* is a best response for player *i* in the perturbed game Γ_i'.

Let

Π_i^* payoff to player *i* from the strategy b^* in the game Γ

$\underline{\Pi}_i$ minimum possible payoff for player *i*, in the game Γ

$\overline{\Pi}_i$ maximum possible payoff for player *i*, in the game Γ

Π_i' maximum possible payoff for player *i* given the strategy b^*_{-i} is used by players -*i* , in the game Γ from any behavioral strategy $b_i' \neq b_i^*$.

Since b_i^* is unique, $\left(\Pi_i^* - \Pi_i'\right)$ and $(\overline{\Pi}_i - \underline{\Pi}_i)$ must be greater than zero.

In the game Γ_i' all the players -*i* all follow their equilibrium strategies b^*_{-i} for the original game Γ. Since the payoffs are not changed in moving from Γ to Γ_i', player *i* cannot receive less than $\underline{\Pi}_i$ and not more than $\overline{\Pi}_i$. With probability $\left(1 - \sum_{h' \in S(h)} e_h^{h'}\right)$ nature selects a path where player *i* makes no mistakes and on this path any strategy $b_i' \neq b_i^*$ gives at most $\Pi_i' < \Pi_i^*$

Thus b_i^* must be a best response to b^*_{-i} in the game Γ_i' if the following condition is satisfied:

$$\left(1 - \sum_{h' \in S(h)} e_h^{h'}\right)\Pi_i^* + \underline{\Pi}_i \sum_{h' \in S(h)} e_h^{h'} > \left(1 - \sum_{h' \in S(h)} e_h^{h'}\right)\Pi_i' + \overline{\Pi}_i \sum_{h' \in S(h)} e_h^{h'}$$

Where the left hand side represents the minimum payoff possible from b_i^*, and the right hand side represents the maximum possible from any other strategy $b_i^{'}$. Since all the probabilities and payoff differentials are non zero, we can re-write this as

$$0 < \frac{\sum_{h' \in S(h)} e_h^{h'}}{\left(1 - \sum_{h' \in S(h)} e_h^{h'}\right)} < \frac{\left(\Pi_i^* - \Pi_i^{'}\right)}{(\overline{\Pi}_i - \underline{\Pi}_i)}$$

Since for any payoffs we can find values of $e_h^{h'}$ such that this is satisfied.

Repeating this argument for all players $i \in N$, we can find values of $e_h^{h'}$,

$$0 < \frac{\sum_{h' \in S(h)} e_h^{h'}}{\left(1 - \sum_{h' \in S(h)} e_h^{h'}\right)} < \min\left[\frac{\left(\Pi_i^* - \Pi_i^{'}\right)}{(\overline{\Pi}_i - \underline{\Pi}_i)} \text{ for all players } i \in N\right]$$

such that all players behavioral strategies b_i^* are best responses. ÿ

(iii). The game Γ contains similar histories and there are two or more equilibria in behavioral strategies. At least one of these will be trembling mind perfect.

Proof: It can be shown that any equilibrium of the original game Γ is for some arbitrarily small value of e_h^h>0 a candidate for trembling mind perfection.

Where $\left(\overline{\Pi}_i - \underline{\Pi}_i\right)$=0, all behavioral strategies for player i are equally good in all possible perturbed games and hence all are trembling mind perfect.

Following a similar argument to that presented in (ii), if $\left(\Pi_i^* - \Pi_i^{'}\right)$>0 then b_i^* can be part of a trembling mind perfect equilibrium.

In the case where $\left(\Pi_i^* - \Pi_i^{'}\right)$=0 the strategy $b_i^{'}$ is a best response to b_{-i}^* (and hence also part of an equilibrium of the game Γ). The strategy b_i^* will be trembling mind perfect unless $b_i^{'}$ is strictly better in all possible perturbed games. If this is the case then $b_i^{'}$ will be trembling mind perfect. ÿ

Bibliography

Abreu, D. and Rubinstein, A. (1988),'The Structure of Nash Equilibria in Repeated Games with Finite Automata', *Econometrica* 56, 1259-1281.

Aumann, R., Hart, S. and Perry, M. (1997a), 'The Absentminded Driver,' *Games and Economic Behavior* 20, 102-116.

Aumann, R., Hart, S. and Perry, M. (1997b), 'The Forgetful Passenger,' *Games and Economic Behavior* 20, 117-120.

Battigalli, P. (1997) Dynamic Consistency and Imperfect Recall,' *Games and Economic Behavior* 20, 31-50.

Binmore, K. (forthcoming) 'A Note on Imperfect Recall,' in *Understanding Strategic Interaction - Essays in Honour of Reinhard Selten*, W.Guth et al. (editors), Springer-Verlag

Gilboa, I. (1997) 'A Comment on the Absent Minded Driver Paradox,' *Games and Economic Behavior* 20, 25-30

Grove, A. and Halpern, J. (1997) 'On the Expected Value of Games with Absentmindedness,' *Games and Economic Behavior* 20, 51-65

Halpern, J. 'On Ambiguities in the Interpretation of Game Trees,' *Games and Economic Behavior* 20, 66-96.

Isbell, J. (1957), 'Finitary Games', *Contributions to the Theory of Games* vol. 3, 79-96 (Annals of Mathematical Studies, 39) (M. Dresher, A.W. Tucker and P. Wolfe, eds.), Princeton: Princeton University Press.

Kreps, D. and Wilson, R.(1982), 'Sequential Equilibria', *Econometrica*, 50, 863-894.

Lipman, B. (1997), 'More Absentmindedness', *Games and Economic Behavior* 20, 97-101.

Myerson, R. B. (1978) Refinements of the Nash Equilibrium Concept', *International Journal of Game Theory* 7, 73-80

Osborne, M. J. and Rubinstein, A. (1994), *A Course in Game Theory*, Cambridge, Mass.: MIT Press.

Piccione, M. and Rubinstein, A (1997), 'On the Interpretations of Decision Problems with Imperfect Recall', *Games and Economic Behavior* 20, 3-24.

Rubinstein, A. (1987), 'The Complexity of Strategies and the Resolution of Conflict: An Introduction', *Theoretical Economics Discussion Paper Series* 87/150, Suntory-Toyota International Centre for Economics and Related Disciplines.

Selten, R. (1978), 'The Chain Store Paradox', *Theory and Decision* 9, 127-159.

Simon, H. A. (1978), 'On how to decide what to do', *Bell Journal of Economics* 9, 494-507.