Putting Trembles Into The Extensive Form *

Tilman Börgers
Department of Economics
University College London
Gower Street
London WC1E 6BT
U.K.

* This paper is a revised and expanded version of some sections of my unpublished paper "On the Definition of Rationalizability in Extensive Games". I am grateful to Martin Hellwig and Joel Sobel for their comments on the earlier paper. I would also like to thank the seminar participants who have listened to presentations of this material, in particular Michael Mandler and Jean-François Mertens who pointed out mistakes in previous versions. The responsibility for remaining errors is mine. I gratefully acknowledge partial financial support from the Swiss National Science Foundation.

Abstract

Selten [7] suggested that one route to a logically consistent theory of extensive games is to assume that players "tremble", i.e. sometimes choose actions at random. This paper describes how these random events can be explicitly included into the extensive form of a game. For this we need to specify in particular the precise points in the sequence of events at which trembles can occur, the probabilities with which they occur, and the information that players hold about trembles. It turns out that the answer to these questions is not obvious. The model that we develop is more general than that of Selten [7]. It includes Selten's model, but also Fudenberg et. al.'s [2] model of normal form trembles as special cases. The paper uncovers implicit assumptions made in previous treatments of trembles. It suggests that these assumptions are more restrictive than necessary. A further issue that the paper addresses is the relation between the strategy space of extensive games that explicitly include trembles, and the strategy space of the underlying games without trembles.

1 Introduction

If a rational economic agent chooses among several possible actions, then his choice will be based on beliefs concerning the outcomes that will result from each of these actions. Hence, in particular, the agent will have some belief about the consequences of irrational actions. If the choice under consideration is part of a game in extensive form, and if further moves of other players follow the choice under consideration, then the consequences of an irrational action will depend, in particular, on the beliefs which other players hold about a player who has been observed to make an irrational choice. A general theory of rationality in noncooperative games must explain how players form such beliefs. The somewhat paradoxical conclusion results that a general theory of rationality in noncooperative games must be based on a theory of irrationality.

Selten [7] introduced a very simple approach. He assumed that rational players cannot avoid "trembles", i.e. accidental mistakes. In Selten's theory these trembles occur with exogenously given probabilities. Selten's approach provides a foundation on which a rational player can base his belief about the consequences of irrational actions. An irrational action by some particular player will be interpreted as a tremble by all the other players, and their behavior following such an action will be based on the belief that the action was a tremble.

Selten's theory of trembles included the important assumption that the fact that a player has trembled once does not make it more or less likely that the player trembles again. More formally, this assumption says that trembles are uncorrelated across different information sets of a player. This assumption has been criticized by Binmore [1] and Reny [4], among others, who referred in this context to Rosenthal's [5] "centipede game." An alternative approach to trembles is due to Fudenberg et. al. [2]. They assume that, before an agent takes any decisions, a random event either determines that this player's behaviour is rational at all information sets of the player, or that this player's behaviour is given by a randomly selected strategy at all information sets. Formally, the assumption is that trembles at different information sets are completely correlated. Trembles of this type are called "normal from trembles." [1] Fudenberg et. al. show that very different results obtain if Selten's trembles are replaced by normal from trembles.

Fudenberg et. al.'s approach to trembles and Selten's approach to trembles clearly form extremes of a whole spectrum of possible theories of trembles. In this paper we undertake a more systematic and complete study of possible models of trembles. We shall follow both Selten and Fudenberg et. al. and regard trembles as random events. Since the extensive form of a game can include random events

[1]Strictly speaking, Fudenberg et. al.'s model describes trembles not as small probabilities that a strategy is played by mistake, but as small probabilities that payoffs are different from what they were in the original game. But an appropriate choice of payoffs can make every strategy dominant. Therefore, I shall ignore this point in this paper.

in the form of "moves of nature", it should clearly be possible to integrate trembles explicitly into the extensive form. Neither Selten nor Fudenberg et. al. have described how this can be done. The first contribution of this paper is therefore to develop an extensive form model of trembles.

The construction of this model brings the treatment of trembles in line with the general postulate of game-theoretic methodology according to which all relevant aspects of a game should be included into the game's extensive form. When specifying an extensive form model of trembles one is forced to be more precise than the existing literature about the order of events when trembles occur, the probabilities with which trembles occur, and the information which players hold about the occurrence of trembles. We show in this paper how these points can be specified in a logically consistent way.

The model of trembles constructed in this paper will include Selten's model and Fudenberg et. al.'s model as special cases, but it will be much more general than either of these two models. Our model thus provides a framework in which the implications of more general assumptions about trembles can be studied.

When describing how Selten's and Fudenberg et. al.'s models of trembles result as special cases of our model, we shall show that both treatments of trembles have several implicit assumptions in common which have not been discussed explicitly before, and which are not necessarily plausible. The two most important such assumptions can be described in words as follows: (i) The probability with which a player chooses an action "by mistake" (i.e. as a result of a tremble) is independent of the action which the player *intended* to choose. (ii) The probability with which different trembles occur does not depend on events which the decision maker does not observe.[2] A brief moment of reflection will show that in reality mistakes may well occur in ways which violate these two assumptions. The model developed in this paper makes it possible to investigate the extent to which analysis can proceed without these assumptions.

An important subject that can be addressed using the framework of this paper concerns the definition of a "strategy." It has been argued that the traditional definition of a "strategy" contains some redundancy in that players have to plan for information sets that cannot be reached given their own planned moves at preceding information sets (Rubinstein [6]). Once the extensive form with trembles has been written out, it becomes clear, however, that a player's plan has to take account of the possibility that the player himself might make mistakes. The supposed redundancies are precisely the contingent plans that need to be made for such cases.

Formally, we shall distinguish two strategy definitions, one being the traditional definition, and another one resulting from the traditional definition if

[2]These statements paraphrase Assumptions 4 and 5 in Section 3 below.

redundancies are removed. It will then be shown that a strategy without redundancies in a game with trembles is the same as a strategy in the traditional sense (with redundancies) in a game without trembles.

This connection between the strategies in games with trembles and in games without trembles is of interest not only in its own right, but also because without such a connection it would not be well-defined to construct the solution of a game without trembles as the "limit" of the solutions of versions of this game with trembles. Such limit constructions were introduced by Selten [7], however he did not investigate in detail the strategy definitions underlying his construction.

The model of trembles that we propose in this paper is only one of several possible models of trembles. An alternative model could, for example, be constructed using an idea of Kreps and Wilson [3]. They argue that, in extensive form games, all moves of nature may be assumed to take place at the beginning of the game, before any personal player has moved. As "trembles" are modelled in this paper as moves of nature, Kreps and Wilson's suggestion is relevant. If one followed their suggestion, the issues discussed below would have to be addressed once one specifies at which time, and to which extent nature's moves are disclosed to the players. The resulting model appears to be more cumbersome than the one presented here.

This paper is organized as follows. Section 2 describes the extensive form model of trembles. Section 3 and 4 show how Selten's and Fudenberg et. al.'s theories of trembles can be regarded as two different special cases of the model of this paper. Section 3 lists assumptions about mistake probabilities that are implicit both in Selten's framework and in Fudenberg et. al.'s model of normal form trembles. Section 4 then shows how Selten's and Fudenberg et. al.'s theories of trembles are special constructions which remain within the framework of Section 3. Section 5 addresses problems concerning the definition of strategies. Section 6 concludes.

2 An Extensive Form Model of Trembles

The model of trembles which we shall introduce in this section will apply to all finite extensive games of perfect recall. But to simplify the exposition we shall assume that the game [3] does not contain any moves of nature, and that there are no "trivial" decision nodes, i.e. decision nodes at which a player has just one possible choice.

As a starting point we recall Selten's ([7], p.35) description of the structure of trembles: "At every information set u there is a small positive probability ε_u for the breakdown of rationality. Whenever rationality breaks down, every choice c

[3]Without trembles.

at u will be selected with some positive probability q_c which may be thought of as determined by some unspecified psychological mechanism."[4]

The information contained in this passage is not sufficient for the construction of an explicit model of trembles. What is missing is whether players' rationality breaks down before or after they make their choice, and whether the concerned player himself and the other players can observe the breakdown of rationality, or whether they observe only the resulting action without knowing how it came about. To obtain an explicit model of trembles we have to specify answers to these questions. In the following, we first summarize the specifications that we choose, then we illustrate them for an example, and then we explain their motivation.

We shall assume that, at every decision node, first the player moving at this node chooses an action, and then "nature" intervenes. With positive probability nature leaves the player's choice unchanged. But, also with positive probability, nature replaces the player's choice by some other choice. All feasible alternative choices have positive probability. Every player observes only those moves of nature that concern himself, but not those that concern other players. Thus, a player is informed if nature replaces his own choice by some other move, and he also learns which alternative move that is, but he does not observe whether the other players' moves are intentional choices or "trembles." Apart from this, players' information is the same as in the original game without trembles.

For completeness we add the assumption that payoffs depend only on the actions actually taken and are identical to the payoffs in the original game. Whether actions were taken intentionally or by mistake hence doesn't affect payoffs. ?

If these assumptions are adopted then, starting from any given extensive game, one can construct a new extensive game which is a version of the original game, but includes trembles. It is easiest to proceed in two steps.[5] First, all considerations of information are ignored. The game tree is expanded, starting from the initial decision node, and following the order of decision nodes in the original game tree. What is being added to the game tree are simply the moves of nature described above. Information sets are then added in a second step. They are constructed in accordance with the assumptions made above. The expansion of game trees is illustrated for Example 1 in Figures 1 and 2.[6] In the following, we shall adopt Selten's [7] terminology, and refer to the game which includes trembles also as a "perturbed" game, and to the original game without trembles as an "unpertubed" game.

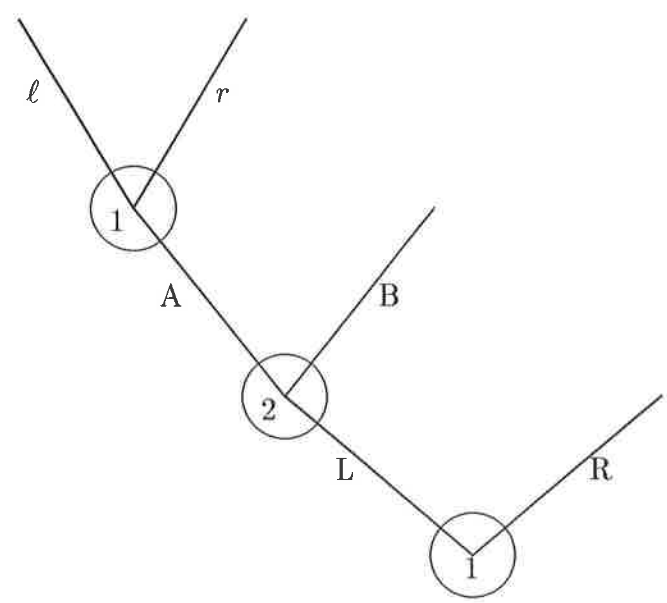
[4]Selten's notation will not be used in this paper.

[5]The procedure that follows is described in order to show that our assumptions are well-defined, and uniquely determine the game with trembles.

[6]In these and all following figures moves of "nature" are represented as moves of "player 0." Also, payoffs are omitted, since they don't matter for the arguments of this paper.

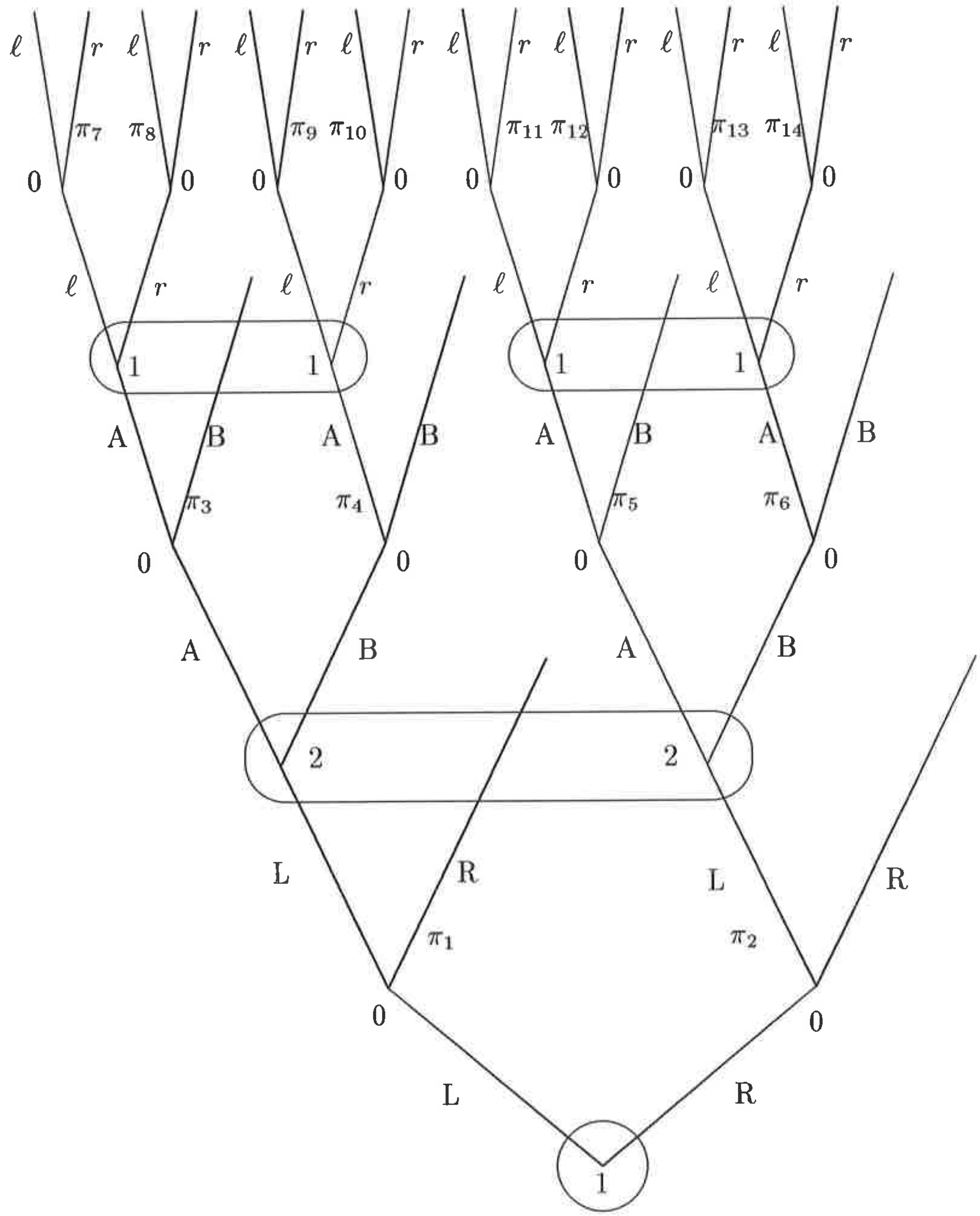
To explain the motivation for the assumptions made, we explain why some alternative assumptions would not have been appropriate for the purposes of this paper. However, as already mentioned in the Introduction, we do not want to argue that the model described here is *the only* sensible model of trembles in the extensive form.

An alternative specification of the order of events is that nature intervenes *before* rather than *after* a player's choice. Specifically, one could assume that nature, with some positive probability, leaves the move to the player, and, with the remaining probability, chooses an action at random. Again, all actions would have positive probability. Unlike our specification, such a model would permit the possibility that "by mistake" a move occurs which the concerned player, if he had been permitted to, would have chosen anyway.



Example 1 (Without Perturbations)

Figure 1



Example 1 (With Perturbations)

Figure 2

Excluding this possibility seems to reflect better the intuitive notion of "mistakes." More importantly, the alternative specification of trembles adopted in this paper would not permit us to employ the conventional notion of a strategy in extensive games, whereas the one which we adopted does have this property. We shall explain in Section 5 how our model relates to the conventional definition of a strategy. To explain the difficulties which the alternative model would create, we consider Example 1. Suppose that player 1 plans to choose L at his initial information set, and now consider player 1 plans for his second choice. If we assumed that nature moves *before* players, then the contingencies for which player 1 would have to plan would include, among others, both the case that he was allowed to choose, and that he decided to choose L, and the case that he was not allowed to choose freely, but that nature chose L for him. This would lead to a definition of strategies which is more complicated than the one conventionally employed.

As regards information, the assumption that players observe all moves of nature which concern their own choices was made because it ensures that the perturbed game is a game of perfect recall. The further assumption that players don't know whether other players' moves were intentional choices or trembles appears to reflect previous authors' implicit understanding of trembles in games. It is also needed to ensure that the conventional definition of a strategy can be used. If this assumption were not made then players could make separate plans for the case that another player makes a particular choice intentionally, and for the case that he makes this choice by mistake. This would lead to a more complicated notion of a strategy than is conventionally employed.

To prepare the arguments of the following sections it is useful to note the relation between a player's information sets in the perturbed game, and the same player's information sets in the unperturbed game. Every information set in the perturbed game is derived from an information set in the original game, and every information set in the original game is mapped into at least one information set in the perturbed game, but there may be several information sets in the perturbed game which correspond to the same information set in the original game. This will be the case if and only if the player choosing at an information set has moved before. Perfect recall implies in this case that in the original game there is exactly one sequence of moves of the concerned player that has preceded the information set. In the perturbed game this will have been the sequence of the player's *actual* moves, but his *intended* moves may have been different. In fact, for any sequence of *intended* moves of the concerned player there is a different information set of the perturbed game. Thus, the information sets of the perturbed games can be identified with pairs consisting of an information set in the not perturbed games, and a history of *intended* actions of the player moving at this information set.

To illustrate this, we consider in Example 1 the second information set of player 1. In the original game this set is preceded by player 1's choice of L in his first decision node. In the perturbed game we have specify in addition which choice player 1 *intended* to make at the initial decision node. This might have been L or R. Hence, there are two information sets in the perturbed game, both of which correspond to the same information set in the unperturbed game.

3 Assumptions for Mistake Probabilities

To describe the probabilities of nature's moves in a perturbed game one needs to specify for every decision node of the game with perturbations, for every action a available at this set, and for every available alternative action $b \neq a$, the probability with which nature replaces a by b if the relevant player chooses a. We shall refer to these probabilities as the "mistake probabilities." The mistake probabilities can be regarded as a complete description of the probabilities of nature's moves because the probabilities with which nature leaves choices unchanged can be deduced as the complementary probabilities.

The purpose of this and the next section is to show how, by making different assumptions about the mistake probabilities in our model we can obtain both Selten's [7] and Fudenberg et. al.'s [2] theories of trembles as special cases. We proceed in two steps. In the current section we present several assumptions for mistake probabilities which are common to both Selten's and Fudenberg et. al.'s theories. In the next section, we then show how the two theories constitute special constructions which satisfy the assumptions of this section. We proceed in this way to point out that some of the assumptions which are common to both previous treatments of trembles are perhaps more restrictive than is desirable.

The first three assumptions in this section, however, are very mild assumptions. Note that in these assumptions expressions like "decision nodes" and "information sets" refer to the game with perturbations, not to the game without perturbations. Assumptions 1 and 2 were already mentioned in Section 2.

> Assumption 1: *At every decision node, for any chosen action, all mistake probabilities are strictly positive.*
>
> Assumption 2: *At every decision node, for any chosen action, the probability that a chosen action is not replaced, i.e. the complementary probability of the mistake probabilities, is strictly positive.*
>
> Assumption 3: *At every decision node, choosing an action will never make it* less *likely that the action is taken.*

Assumptions 1 and 2 are simple "full support" assumptions. If Assumption 3 were not satisfied, it would not be clear what it would mean to "choose an action." Assumption 3 uses the expression "not less likely" rather than the expression "more likely" in order to make the model compatible with normal form trembles. This will become clear in Section 4 below.

To illustrate these assumptions consider again the perturbed game in Figure 2. In this example the vector of mistake probabilities, π, has 14 entries: $\pi = (\pi_1, \ldots, \pi_{14})$. We have indicated these probabilities in the figure. The vector π satisfies Assumptions 1 and 2 if all of its components are contained in the open interval (0,1). Assumption 3 is satisfied if in addition $1 - \pi_k \geq \pi_{k+1}$ holds for all odd $k \in \{1, \ldots, 14\}$.

The next two assumptions are more restrictive.

> Assumption 4: *At every decision node the mistake probabilities are independent of the action chosen, i.e. the probability with which a chosen action* a *is replaced by an action* c *is equal to the probability with which a chosen action* b *is replaced by* c.
>
> Assumption 5: *If two decision nodes belong to the same information set then the mistake probabilities at these two nodes are identical.*

Intuitively, Assumption 4 means that the probability that a player makes any particular mistake does not depend on the action that the player *intends* to take. Assumption 5 says that the probability with which a player makes a certain mistake can depend on a particular event only if the player can observe this event. Whatever a player doesn't know, cannot influence this player's mistake probabilities. Both of these assumptions are in conflict with my introspection. Therefore, I regard these assumptions as more restrictive than Assumptions 1-3.

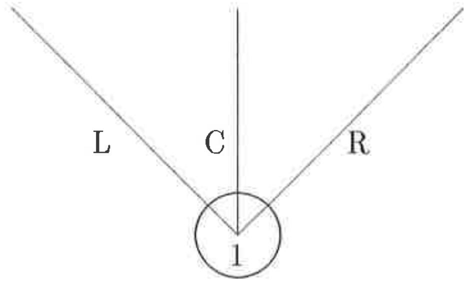
Assumption 4 is void in Example 1. To illustrate the meaning of this assumption we therefore introduce Example 2, a one player game which is shown in Figures 3 and 4.

Figures 3 and 4 (See next page.)

In this example, Assumption 4 requires $\pi_3 = \pi_5$, $\pi_1 = \pi_6$, and $\pi_2 = \pi_4$.

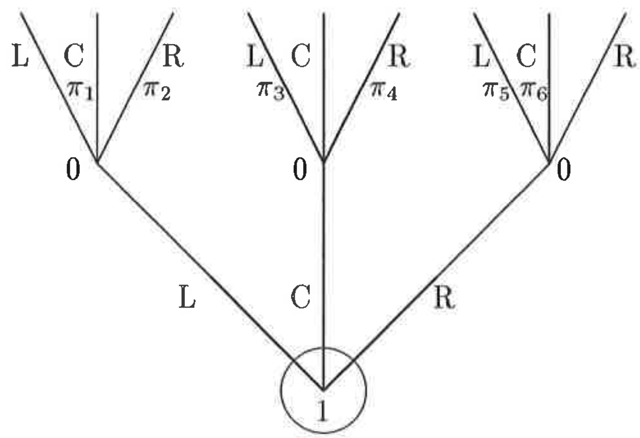
Assumption 5 is void in Example 2, but it can be illustrated by means of Example 1. There, Assumption 5 requires that $\pi_k = \pi_{k+2}$ for k =3,4,7,8,11 and 12.

The effect of Assumptions 4 and 5 together is that for every information set in the perturbed game, and for every action that is feasible at that information set, there is a unique, well-defined probability that this action is taken by mistake. In the following, we shall call this probability the "mistake probability for the action at the information set." The simplicifcation achieved by



Example 2 (Without Perturbations)

Figure 3



Example 2 (With Perturbations)

Figure 4

Assumptions 4 and 5 is that we don't have to condition the mistake probabilities on the *intended* actions, or on the decision node.

Assumptions 1 and 2 ensure that all mistake probabilities are strictly between zero and one. Moreover, as one can easily see, Assumptions 1, 2, 4 and 5 imply that Assumption 3 holds if and only if the sum of the mistake probabilities at any given information is at most one. [7]

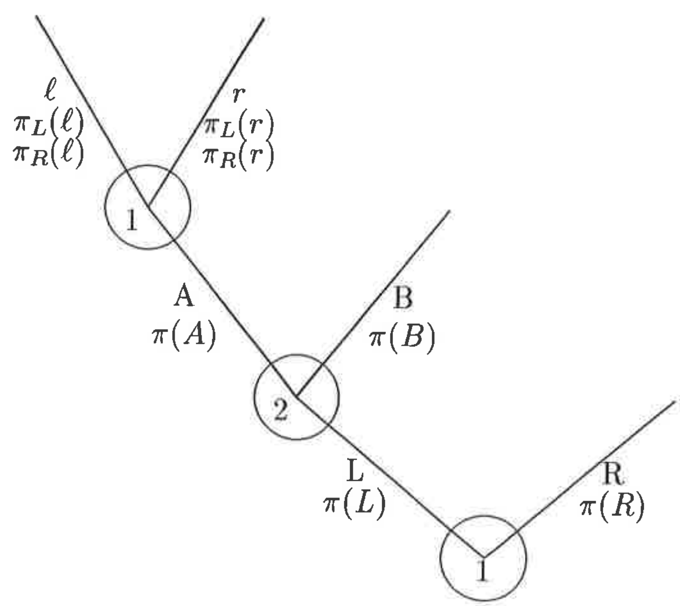
Recall from Section 2 that the information sets of the perturbed games can be identified with pairs consisting of an information set in the not perturbed games,

[7]Proof: Assumption 3 requires that for any information set and for any action a available at that information set: (Mistake probability for a) $\leq 1 -$ (Sum of mistake probabilities for actions other than a) $\Leftrightarrow$ (Sum of all mistake probabilities) ≤ 1.

and a history of intended actions of the player moving at this information set. If Assumptions 1-5 are imposed we can hence associate with each such pair a vector of mistake probabilities that all lie between zero and one, and that add up to at most one.

We illustrate this for Example 1. We also introduce some further notation for this example. This notation will be useful when we return to Example 1 in the next section. For the first two information sets in Example 1 in Figure 1 there is just one corresponding information set in the perturbed game in Figure 2. We shall denote the mistake probability for action L (resp. R) at the first information set by $\pi(L)$ (resp. $\pi(R)$), and we shall denote the mistake probability for action A (resp. B) at the second information set by $\pi(A)$ (resp. $\pi(B)$). There are two information sets of the perturbed game which are derived from the third information set of the unperturbed game. They correspond to the two histories "L" and "R"[8]. We define $\pi_L(\ell)$ to be the mistake probability for action ℓ at the information set associated with history L, and we define analogously $\pi_L(r)$, $\pi_R(\ell)$ and $\pi_R(r)$.

This notation is summarized in Figure 5:



Example 1 (Under Assumptions 1-5)

Figure 5

Assumptions 1 and 2 together require that all probabilities defined in the preceding paragraph are in the open interval (0,1). Assumption 3 is satisfied

[8]The histories have thus just one component.

at any point in the course of the game, the probability that some mistake occurs in the remainder of the game is low. The second definition is more restrictive.

In Example 1, as represented in Figure 2, the first definition requires π_k to be close to zero for all k with $1 \leq k \leq 10$. The second definition requires π_k to be close to zero for all k with $1 \leq k \leq 14$.

If we impose Selten's assumption then the two definitions are equivalent. If we make the assumptions underlying normal form trembles, and if players move repeatedly, then we shall not be able to satisfy the second definition. This is because Assumption 8 contradicts what is required by this definition. The first definition will be satisfied if we choose the minimum probabilities of normal form strategies to be small.

But the two definitions presented above are well-defined independent of the restrictive assumptions underlying the Selten's model of trembles, or the normal form model. Moreover, it appears that one can obtain interesting results from these assumptions in conjunction with the rather mild Assumptions 1-3, thus avoiding the more restrictive Assumptions 4-8. One can thus hope that the distinction between two definitions of "small" mistakes can supersede the distinction between the two models of trembles which the literature has considered.

References

[1] Binmore, K.G., Modeling Rational Players I, *Economics and Philosophy* 3 (1987), 179-214.

[2] Fudenberg, D., D. Kreps and D. Levine, On the Robustness of Equilibrium Refinements, *Journal of Economic Theory* 44 (1988), 354-380.

[3] Kreps, D. and R. Wilson, Sequential Equilibrium, *Econometrica* 50 (1982), 863-894.

[4] Reny, P., Common Belief and the Theory of Games with Perfect Information, *Journal of Economic Theory* 59 (1993), 257-274.

[5] Rosenthal, R.W., Games of Perfect Information, Predatory Pricing, and the Chain Store Paradox, *Journal of Economic Theory* 25 (1981), 92-100.

[6] Rubinstein, A., Comments on the Interpretation of Game Theory, *Econometrica* 59 (1991), 909-924.

[7] Selten, R., Re-Examination of the Perfectness Concept for Equilibrium Points in Extensive Games, *International Journal of Game Theory* 4 (1975), 25-55.