



UNIVERSITA' DEGLI STUDI DI TRENTO - DIPARTIMENTO DI ECONOMIA

---

# **A Fuzzy Logic and Default Reasoning Model of Social Norm and Equilibrium Selection in Games under Unforeseen Contingencies**

**Lorenzo Sacconi  
And  
Stefano Moretti**

---

Discussion Paper No. 12, 2004

The Discussion Paper series provides a means for circulating preliminary research results by staff of or visitors to the Department. Its purpose is to stimulate discussion prior to the publication of papers.

Requests for copies of Discussion Papers and address changes should be sent to:

Dott. Stefano Comino  
Dipartimento di Economia  
Università degli Studi  
Via Inama 5  
38100 TRENTO ITALY

# A Fuzzy Logic and Default Reasoning Model of Social Norm and Equilibrium Selection in Games under Unforeseen Contingencies

Lorenzo Sacconi<sup>1</sup> and Stefano Moretti<sup>2</sup>

<sup>1</sup> Department of Economics, University of Trento and EconomEtica, Interuniversity centre of research, Bicocca University, Milano

<sup>2</sup> Department of Mathematics, University of Genova  
and Department of Environmental Epidemiology, National Cancer  
Research Institute of Genova, Italy.

## Abstract

This paper focuses on the role that social norms play in the selection of equilibrium points seen as social conventions under unforeseen contingencies – that is, their role in the emergence of regularities of behavior which are self-enforcing and effectively adhered to by bounded rational agents due to their self-policing incentives. Differently stated, given a set of game situations imperfectly described, we want to understand how general and abstract norms provide at least the starting point for a norm-based equilibrium selection reasoning procedure which in the end will be able to determine which equilibrium point, belonging to perfectly described games, will be played as the unique solution of each imperfectly described game. In order to solve such a problem we introduces a selection process based on the reformulation of default logic in terms of possibility theory.<sup>1</sup>

**Keywords:** Social Norms, Equilibrium selection, Unforeseen Contingencies, Fuzzy Logic, Possibility Theory, Default Reasoning.

---

<sup>1</sup> A shortened version of this paper has been presented at the NMR 04 International Workshop on Non Monotonic Reasoning, 6-8 June, Whistler BC, Canada

## 1. Introduction and motivation

We are not concerned here with the explanation that norms induce equilibrium points in games because of their structuring effects on payoffs and incentives. Equilibrium points are seen as pre-existing possible combinations of strategies in a given game situation. Norms, on the other hand, are primarily seen as sets of logically consistent normative statements prescribing a single action for each player in any game situation of a given class of games. Hence, these statements simply ask players to adopt a given equilibrium behaviour within the existing equilibrium set of the given game situation. Their effectiveness does reside only in their functioning as cognitive devices which induce the appropriate system of expectations about equilibrium behaviours. In fact, given the appropriate system of expectations, players have endogenous incentives to adopt the prescribed behaviour, as long as they expect the other players to execute the corresponding actions.

Note, however, that most non-trivial non-cooperative games have multiple equilibrium points. Mixed-interest coordination games or the typical iterated Prisoners' Dilemma, where the Folk theorem applies, can be regarded as the natural classes of game situations to which norms refer. Consequently, given multiple equilibria, there are also multiple possible norms in the meaning just defined. The multiplicity problem then ensues: given a multiplicity of equilibria - such that the expectation of each equilibrium supports the decision to comply with a different norm in turn - which norm should be chosen or followed? We assume that an equilibrium point gains its prescriptive force in terms of individual rational choice since it is common knowledge<sup>2</sup> that the equilibrium itself will be adopted as the solution of a game, i.e. it has gained general acceptance. Hence, multiple equilibria are situations such that, even if we know that some equilibrium point will be adopted, there is no implication as to which equilibrium will be adopted, and as a consequence we have no reason to act according to any of the possible equilibria. Unless we know which norm is to be adopted in order to solve the game, we cannot say that we have the belief system concerning the equilibrium behaviour that grounds a given norm in terms of the incentive-based reasons to act upon it.

What we are looking for relates to the explanation that equilibrium points are selected because they are *focal points* or exhibit what is usually called *salience* (Shelling 1960, Lewis 1969). Most interesting approaches, however, view focal points not as primitive brute psychological facts attached to behaviors, but rather as the result of existing social norms (Kreps 1990). Hence norms work as the determinants of the state of knowledge that makes an equilibrium the predictable outcomes of the game, while the equilibrium provides in its turn endogenous incentives to comply with the relevant norm. There is, however, a sort of circle here: equilibrium points are selected because of focality, which in turn is due to social norms that give "salience" to that equilibrium. But on the other hand this may be so only because some of the possible multiple norms are focal points on their own account, ones able to coordinate expectations in the relevant class of games. Which requires no less explanation than the salience of equilibrium points as such.

The relevant question to ask is therefore the following: where does the basic shared knowledge about norm-following behaviour come from? There must be an independent

---

<sup>2</sup> For more information on the definition and use of the concept of *knowledge* see for instance Aumann (1976), Binmore & Brandenburger (1990), Fagin et al. (1995), Brafman & Tennenholtz (1997).

process of reasoning which induces us to believe that everybody will comply with the norm during the given game in order to generate the system of expectations which rationally justifies a player's adoption of his equilibrium strategy because of its best response property. We shall seek to deal with this problem by investigating how players come to believe that in a given game all the players will conform with a general norm before they begin calculating that, because it is known that all of the players will conform with the norm, they then play the given equilibrium. To simplify, let us introduce an *ex ante/ex post* distinction into players' reasoning processes: it is *ex post* reasoning to say that, because it is common knowledge that a given equilibrium is the solution of the game, then players have the incentive to comply with it. By contrast, it is an *ex ante* reasoning what takes us from the knowledge of a multiplicity of pre-existing norms, each possibly endowed with prescriptive meaning as far as it is generally accepted as the solution concept appropriate for a given game class, to the conclusion that one of these norm is in fact the one also accepted as the solution theory in the given game that concerns us, so that the *ex post* common knowledge-based reasoning may then emerge. By 'equilibrium selection process' is meant the *ex ante* reasoning that brings the players to a state of shared knowledge that a given equilibrium will be the solution of the given game at hand.

Equilibrium selection is a topic widely addressed in game theory (Harsanyi 1975, Harsanyi-Selten 1988, Binmore 1987/88, Fudenberg-Levine 1998, Samuelson 1997). Harsanyi and Selten put forward the first theory of equilibrium selection. They ground it on a model of the outguessing regress – which consists of the supposition that players reciprocally simulate the each others' reasoning processes by repeatedly assigning their own previous reasoning steps to the counterpart – an approach that Binmore (1987) calls “*eductive*” (and which we shall adopt in this work). But, they developed their theory under overly strong epistemic assumptions - so strong that not even their equilibrium selection model could satisfy them (Sacconi 1995b).

However, we are looking for a solution to the equilibrium selection problem which has not been addressed to until now. It ranges over games where equilibrium points must be selected by players with incomplete knowledge of the possible states of the world, so that players cannot say without ambiguity the game in which they are involved. Multiplicity occurs in this context because of the ambiguity concerning the game that players are about to play. This can be construed as if they were able to predict the single equilibrium with which they could solve a game if there were no ambiguity about the game that they are going to play. Nevertheless, multiplicity returns because of the ambiguity about the game that they are to play occasioned by unforeseen contingencies. Note that this is not the same as extending the standard approach of games under incomplete information (Harsanyi 1967/68). It is not the case that players are uncertain about the game selected by Nature, because they do not know the occurring state of the world. They perfectly know to be in an unforeseen state. The problem is that in this state the game to be played is imperfectly described – i.e. it is ambiguous what norm must be applied to the game under consideration. Our approach is to model how players come to the conclusion that a given ambiguous game falls within the domain of a pre-existing general norm. The point to be understood is how players come to recognize that a given game situation falls in the domain of a general abstract norm so that there are practical reasons for them to comply with that norm, while everybody else must also follow the same norm.

In a sense this is not to eliminate the presumption that there exists some “*common knowledge*” in the world about how players will act given certain norms and forms of

interaction. This presumption can be retained, but we want to minimise its impact. In fact, we break the implication that because we know that a general norm is followed in some games class, then every player knows that in the particular game that they are playing (which belongs to the class) the norm will be followed. The chain of reasoning is broken where it states that because there is a norm, and because it is shared knowledge that it is followed in a certain class of game situations, then it is also shared knowledge that it will be followed in our game. The assumption within brackets is not innocuous and must be questioned. This inference cannot be made as a matter of pure logic, nor can it be taken for granted as a primitive psychological fact. It must be explained by explicit modelling of the approximate reasoning process performed by the players when they face game situations and reason as to whether these games can be understood according to the pattern of a solution theory devised for a general abstract game form. It is by this reasoning that we can explain from whence derives the required “common knowledge” (or something less than common but which can perform the same role - call it “shared knowledge”) in the concrete game situations that the players face.

We intend to show that such reasoning process employs fuzzy logic and default reasoning<sup>2</sup>. Fuzzy logic is used in order to represent the cognitive situation of ambiguity about the game to be played and how the player uses social norms as a device of pattern recognition of unforeseen contingencies. Default reasoning seems appropriate because each player under this situation of ambiguity does not hold a completely sound theory of the behavior of his opponent. Nevertheless, as far as this is his best information about how a rational player will act in an unforeseen situation, he may use his own model of behavior in order to simulate those of the opponent. Given that, unless an evidence of the contrary emerges, he repeatedly attributes his own model of norm-based deliberation to the other player, equilibrium selection results as a reasoning process subsuming new games under existing general norms by tentatively extending their domain of application.

The paper proceeds as follows. First (sec.2), we define two versions of a basic game such that in each of them it is assumed that players know that a given solution (norm) is shared knowledge (the Nash product for a “demand game” and the typical non-cooperative solution for an “enlarged DP” game – both coinciding with equilibrium points of two versions of the same basic game). Then (sects. 3 and 4), we introduce the basic incomplete knowledge situation by assuming that a move of Nature selects unforeseen states of world where one of the two versions of the basic game could be played. Default reasoning then enters the scene. In section 5 we introduce a proper selection process based on the reformulation of default logic in terms of possibility theory suggested - after Zadeh (1978) - by Dubois, Prade and Benferhat (Dubois& Prade (1995a-b-1996), Dubois et al. (1992)) together with many applications of possibility theory in decision theory (Dubois et al. (1997), Dubois et al. (1999)). We end by suggesting that the resulting equilibrium will be supported by what in default logic is called an “extension” of a given theory - which is usually characterised as a fixed point (Reiter 1980) - obtained by the iterate application onto itself of the set of accepted defaults, without introducing any contradiction. A remarkable feature of the result is that the equilibrium achieved rests essentially on the non-monotonicity of default logic, and hence

---

<sup>2</sup> As far as we know, the first time that default logic was suggested as one possible approach to the foundations of some game theoretical concepts was Bacharach (1994) with reference to rationalisability. However, here we complement it with fuzzy logic and possibility theory and address a different game theoretical problem.

allows for mistakes and revisions which seem to pertain to the very nature of bounded rationality.

## 2. Two games of reference

The first game, G1, is the typical demand game. 100 dollars are to be divided into two shares according to demands made by two players. If shares sum to 100 or less, the players obtain what they demand, but if the demanded shares sum to more than 100, they do not obtain anything. The players participate in the game by announcing proposals to share that they accept and that they think can be compatible with the proposal to share that the other player is simultaneously making. In fact, they make demands and offers at the same time, and only once. They can also refuse to demand or offer anything, and this will end the game notwithstanding the other player's proposal, even though he would have been ready to concede the entire sum to the first player. If one or both refuse to agree, the result is the *status quo* (zero dollars). Thus players must guess demands that may be acceptable to the counterpart, under the condition that only couples of demands that sum to no more than 100 are allowed.

Let us assume (this is where the announced assumption of background shared knowledge comes in) that the players know that in games like this there is a unique rational solution: that is, maximising Nash bargaining function (N.b.f.) which, under the appropriate utility representation, coincides with the 50-50 sharing rule. We simply assume that it is shared knowledge that in this game players will adopt the Nash bargaining solution because they accept the bargaining axioms given by John Nash (or the equivalent Harsanyi-Zeuthen's axiomatisation). This is an abstract norm generally accepted for solving games like the one under consideration, and it is shared knowledge that it is so. It also requires one more condition: games like the demand game are *cooperative bargaining games*; which means that the demands and offers made by a player are binding for him and, as far as they are accepted by the counterpart, the proponent cannot renege on them - i.e. he must comply with any agreement once it has been reached. Let us consider a situation in which two compatible proposals are made. A pair of demands and offers determining an agreement then results, and it will necessarily be executed. The player cannot make defection (in the remaining of the paper strategy D) on the agreement: he cannot refuse to concede the share offered and cannot expropriate the full sum by cheating the other party to the agreement. This amounts to saying that, within the demand game, it is shared knowledge that the game is *cooperative*, which implies that one more rule - *pacta sunt servanda* - applies to it, and it is shared knowledge that it is binding. There may be many reasons why it is known that this is an effective rule of the game (it may be that neither player is physically able to breach announced agreements, or the players may dislike cheating in the sense that the payoffs which they attach to defection are zero, or the game could be embedded in an un-modelled and even larger repeated game, where respecting pacts would be a self-enforcing convention). Here we do not ask why this is so, and simply assume that inasmuch as the players understand the demand game - so that it is shared knowledge that N.b.f. is the accepted solution - they must also understand that players cannot successfully make defection. The main assumption in game G1 is that players accept the cooperative theory of bargaining (a set of normative statements): this is shared knowledge. We may say that games like G1 belong to the domain of a general abstract norm: the solution theory  $\sigma_1$ , i.e. N.b.f.

Next, let us assume that the players also know that there are some other games where this is not the case. As far as the opportunity to share 100 dollars by agreement is concerned, these further games are completely identical to the demand games. However they are *non-co-operative* – i.e. they are game situations that do not satisfy the requirement that all players accept the rule *pacta sunt servanda*.

This is the case of the second game G2 consisting of a slight payoff change to the demand game G1. In G2 the defection strategy D has an effective role to play, because the players are now able to free-ride the counterpart by refusing to co-operate, and at the same time by appropriating the entire sum if the counterpart decides to make a proposal for agreement. This can be read as cheating the other player and obtaining the entire sum after the other player has made a proposal. For example, let the second player make any proposal. What he immediately sees is the first player refusing to co-operate by leaving the bargaining table and tipping it over. Next, however, the first player considers the situation more carefully and realises that the entire sum to be shared has been stolen. According to the rule of the demand game, that sum should be forgone by both players. On the contrary, in game G2 the defecting player may obtain the entire sum by taking unilateral advantage of the other player's disposition to agree. There is no assumption of any co-operative structure underlying this game. Moreover, G2 is meant to represent that no agreement has any chance of being self-enforced because the defection strategy is dominant in the game (which is in fact an enlarged PD game). The explicit representation through the defection strategy of the ability to frustrate any cooperative endeavour and gain advantage from this conduct shows that the underlying game situation no longer belongs to the domain of cooperative game theory. *Pacta* are not seen by the players as being *servanda*, and there is shared knowledge that the rules of the game do not imply that any agreement will be enforced by some unrepresented mechanism on the back of the situation. Moreover there is no shared knowledge that N.b.f. is the accepted norm in this game. On the contrary, there is a quite different norm accepted by both the players, and whose acceptance by them is shared knowledge: the solution theory  $\sigma_2$ , i.e. "if there is an outcome in dominant strategies, it can be expected to be the solution of the game".

### **3. Unforeseen states of the world: vagueness of the appropriate solution concept**

The foregoing section was only preparatory to our main analysis. Consider the following game situation. Two players P1 and P2 are in a position to play a game under incomplete knowledge, which means that they can be subjected to choices by Nature that selects states of the world which they were not yet able to imagine before the start of the game. We assume that they know the two abstract classes of games - the demand game and the enlarged DP game - depicted above. They have the necessary background in game theory, and consequently also know the accepted solution theories to be applied to the two prototypical game situations, i.e. the two solution theories  $\sigma_1$  and  $\sigma_2$  that, according to our point of view, are two established social norms. They of course have shared knowledge that if the game that Nature induces them to play were G1, it would be obvious for them to play according to the norm  $\sigma_1$  (N.b.f.), but if Nature were to involve them in a G2 game then it would be obvious to them that they must adhere to the norm  $\sigma_2$ . That is to say, at the back of their minds P1 and P2 have the representation of a set of possible states of the world which all and only match one of two possible events - having to play G1 or G2 - and they expect to be



confronted by Nature's choices that will put them in states where alternatively G1 or G2 are to be played.

The game starts with a Nature's move. Nature may select one of two states  $w_1$  and  $w_2$ , in each of which it is true that either G1 or G2 (but not both) will be played. However, Nature may also select (or, more exactly, let the players discover) a set  $\Omega$  of unforeseen (by the players) states of the world, which are states that do not conform with the properties and predicates in use in P1's and P2's language with which they in general describe games. Owing to their unforeseen characteristics, these states do not exactly match the foreseen events that the players are prepared to learn. That is to say, in each  $\omega \in \Omega$  it is *neither* true *nor* false that G1 or G2 is to be played. Note that this is not the usual game under uncertainty: it is not the case that the players have information partitions defined over the possible states' set which are not refined enough to discriminate exactly the occurrence of states where it is true that the game G1 is to be played from the occurrence of states where on the contrary it is true that game G2 is to be played. After Nature's choice, the players discriminate perfectly among the states and there is no uncertainty about which state has been reached. The situation is meant on the contrary to represent the ambiguity or vagueness that the players face, as long as *any* state is scrutinised, about whether in that particular state the game G1 or G2 is to be played. Vagueness is the consequence of their lack of *ex ante* knowledge about the entire set of states that *ex post* Nature reveals to be possible, which implies that these are unforeseen states. Thus vagueness is the consequence of the players' lack of the conceptual and linguistic tools with which to clearly describe each  $\omega \in \Omega$  in any detail or, in other words, the consequence of the fact that each  $\omega \in \Omega$  is unclear as to the game to be played in it. If it were clearly specified *ex post* whether in any state  $\omega \in \Omega$  the game G1 or G2 is to be played, then the players would have been able to foresee these states at least as states of the world that *ex ante* they knew to be possible.

According to one of the present authors (Sacconi 2000, 2004), however, unforeseen states of the world can be analysed in terms of fuzzy events, which are the fuzzy sub-sets of the reference set  $\Omega$ . A general norm, construed as the solution for a given abstract class of games, has a domain of application (a set) which is defined by the membership functions of each state in the set representing that domain. Foreseen states have crisp membership functions (from the beginning we are able to say whether these states, if they occurred, would make a case for a specific norm). Unforeseen states, by contrast, cannot *a priori* be classified as belonging to one or another. When they are revealed the judgment about their belonging results to be vague. Hence, *ex post* they have fuzzy (graded) membership functions taking their values into the real line  $[0,1]$  and expressing the degree to which the *ex post* revealed states can be recognised as belonging to the domain of any given norm.

So far, we have had two possible norms, each prescribing a unique equilibrium point of the game within which it is the accepted solution theory. But in the game under incomplete knowledge and unforeseen states of the world, players cannot say at first glance which equilibrium point will come about, because they cannot say which solution theory is appropriate to the game that they are going to play (or, better, in the current state of the world). We want to understand how general and abstract norms provide at least the starting point for a norm-based equilibrium selection reasoning procedure which in the end will be able to determine which equilibrium point, belonging either to game G1 or game G2, will be played as the unique solution of the game. As expected, the answer will prove to depend on

Nature's choice of the state of the world. This does not introduce any uncertainty, nor does it make room for any probabilistic nature of the equilibrium selection procedure. Nature first makes its choice but then players learn it without any uncertainty before they reach their own move in the game. If Nature did not select unforeseen states of the world, players P1 and P2 would learn with no ambiguity whether they were going to play G1 or G2, and they would solve each game according to its accepted solution theory. Consequently, we need only investigate what happens when any  $\omega \in \Omega$  is chosen by Nature.

#### 4. Solution domains as fuzzy sets

Let us start by assuming that Nature has two consecutive moves at the beginning of the game. First it chooses among  $w_1$ ,  $w_2$  or  $\Omega$ . Under  $w_1$  and  $w_2$  it is true that games G1 or G2 are played respectively. We disregard these branches of the game tree because there is a completely obvious solution for each of them. Next, Nature has its second move among states in  $\Omega$ .

Take  $\Omega$  to be the set of all the possible, alternative, complete descriptions of the interaction situations between players P1 and P2, which can be worked out by affirmation or negation of any concrete property that may influence the structure of their interaction - expressible by predicates in the language that will evolve in order to describe the unexpected features of  $\Omega$ . Assume that any  $\omega \in \Omega$  typically specifies the characteristics of the players, the principles of ethics that they follow, the set of options open to them, the environment that surrounds them, and every other property that makes it possible to say whether the norm *pacta sunt servanda* holds or does not. Thus,  $\Omega$  is the "universe of discourse" containing all the possible concrete game situations that the players may be involved in once Nature has made its second choice. Remember, however, that some of these concrete characteristics are genuinely "unexpected" and are not univocally expressible by means of terms belonging to the *ex ante* language which was in use when the description of the two games G1 and G2 was under consideration. This means in fact that, within the *ex ante* in use language, some predicates cannot be exactly affirmed or denied when the states of the world  $\omega \in \Omega$  are considered. In particular, under each state  $\omega \in \Omega$ , some of their describable characteristics cannot be traced back unambiguously to the conditions established as necessary to state whether the game form G2 or the game form G1 is the appropriate abstraction from the concrete game situations under consideration. Thus, it is not clear whether possession of such characteristics or their negation implies that game G1 or game G2 will be played.

Then, let  $\Omega$  be the set of *ex ante* unforeseen states that Nature may select at its second move. Hence define two fuzzy events

$$\underline{G1} = \{\omega, \mu_{\underline{G1}}(\omega) \mid \forall \omega \in \Omega\} \text{ and } \underline{G2} = \{\omega, \mu_{\underline{G2}}(\omega) \mid \forall \omega \in \Omega\}.$$

The fuzzy membership functions  $\mu_{\underline{G1}}$  and  $\mu_{\underline{G2}}$  represent the vagueness associated with the occurrence of each state concerning whether the game to be played is G1 or G2. As  $\underline{G1}$  and  $\underline{G2}$  are sets of states defined by associating each unforeseen state  $\omega_i$  with its membership function in the relevant set, they may be understood as *fuzzy events* that occur when unforeseen states are selected. These sets may also be understood as the vague domains of application of the two solution theories  $\sigma_1$  and  $\sigma_2$  respectively (for example  $\sigma_1$  applies to  $\omega_1$  at degree 0.9).

A usual, we can consider any fuzzy event as a flexible constraint on the *possibility* of states, according to the expression

$$\pi_{X \text{ is } \underline{G}_j}(\omega) = \mu_{\underline{G}_j}(\omega), \quad \forall \omega \in \Omega, \forall j=1,2$$

Let  $X$  be a variable ranging over the set of states  $\Omega$  constrained by the fuzzy information that “ $X$  is  $\underline{G}_j$ ”: that is, the game played in the state selected out of  $\Omega$  is  $\underline{G}_j$ ,  $j=1,2$ , according to the relevant membership function. Saying that  $X$  is a variable ranging over the universe of discourse  $\Omega$  means that  $X$  takes its values from the set of concrete descriptions of the possible interaction situations between  $P1$  and  $P2$ , which is in fact what any state in the set  $\Omega$  describes. The players’ information constrains the possibility that  $X$  takes any value in  $\Omega$ : that is, the admitted values of  $X$  are restricted to those that are compatible with the information that any given  $\omega \in \Omega$  belongs to  $G1$  or  $G2$ . Given the unexpected nature of these states, however, variable  $X$  is constrained by fuzzy information about the identification of any interaction situation with the game form  $G1$  ( $G2$  respectively). Thus the possibility of any state  $\omega \in \Omega$ - given the fuzzy information about whether the game played under  $\omega$  belongs to  $\underline{G}_j$  - sums to the degree of the state’s belonging to  $\underline{G}_j$ .

### 5. A possibilistic account of the equilibrium selection default reasoning

Items of incomplete knowledge may be described by possibilistic logic formulas, i.e. first-order logic formulas with a numerical weight between 0 and 1, which is a lower bound on a possibility measure  $\Pi$  (Dubois&Prade 1996). Thus, in the following our model is developed in terms of a possibilistic modelling of default reasoning. Main reference for that are (Benfherat et al. 1992; Benfherat et al. 1997; Dubois&Prade 1995a; Dubois&Prade 1996).

We here use this possibility approach to the study of the default equilibrium selection reasoning in a game situation where Nature selects unforeseen states of the world, so that it is vague which of two games  $G1$  and  $G2$  has to be played. We confine our analysis to only two unforeseen states  $\omega_1$  and  $\omega_2$ . In order to keep thing as simple as possible, we also simplify the representation of the two games, skipping all the strategies but those coinciding with solution concepts. Thus consider the two simplified games, where  $G1$  is (in the following games, players’ utility functions are taken to be identical to linear transformations of the monetary payoffs mapping money onto the real interval  $[0,1]$ )

<b>P1 \ P2</b>	<b>50</b>	<b>D</b>
<b>50</b>	0.5 0.5	0 0
<b>D</b>	0 0	0 0

In such a game there is shared knowledge that  $\sigma_1$  is the appropriate solution theory which requires to play the pure strategy 50, whereas game  $G2$  is

<b>P1 \ P2</b>	<b>50</b>	<b>D</b>
<b>50</b>	0.5 0.5	0 1
<b>D</b>	1 0	0.21 0.21

where there is shared knowledge that  $\sigma_2$  is the accepted solution concept which requires to play the pure strategy  $D$ . Hence the relevant events, meant as domains of the two solution

concepts  $\sigma_1$  and  $\sigma_2$ , are the two fuzzy sets  $\underline{G}_1 = \{(\omega_1, 0.5), (\omega_2, 0.2)\}$  and  $\underline{G}_2 = \{(\omega_1, 0.4), (\omega_2, 0.4)\}$ . Next, we will describe the players' reasoning in four steps:

- *Step 1*: we start by considering the possibilistic expected utility calculation that each player performs given his information on the two prototypical game  $G_1$  and  $G_2$  and the fuzzy knowledge on states  $\omega_1$  and  $\omega_2$ , in order to derive his first conjecture on the solution strategy to be used;
- *Step 2*: the second step in player's reasoning process is his attempt to guess the first step in the reasoning process just executed by the counterpart;
- *Step 3*: each player may now check whether the conclusions that he reached in the first step, relative to the strategy to be used in each unforeseen state  $\omega_1$  and  $\omega_2$ , are stable against the new beliefs that he has gained by default about the other player's action in each state;
- *Step 4*: players iterate the previous steps. It turns out that if by analogous default rules player  $P_1$ 's reasoning at step 3 is assigned to player  $P_2$ , in order to recalculate his strategy choice at the third step of his reasoning process, player  $P_2$  can only be predicted to reach the same conclusions as player  $P_1$ .

### **5.1. First step in the equilibrium selection process: calculating the first hypothetical best response of a player**

To begin with, comment is required about the form of the players' utility functions. These take the form of fuzzy utility functions understood as measures of the degree to which a decision satisfies a flexible constraint in terms of intermediary levels of capability (Yager 1979). Let the constraint be expressed by the following requirement: "take the largest part of 100 dollars under the feasibility condition that shares gained by the two players must sum to no more than 100". Utility functions express the rank of capability associated with any decision, conditionally on the other player's decision, to satisfy the given constraint. This is expressed by means of a membership function which maps the decision set on the real line  $[0,1]$  representing levels of satisfaction. Hence, utility values coincide with membership degrees of a fuzzy set defined over the strategy set of each player taken as the typical reference set. These utility functions have the form appropriate to enable us carrying out the following exercise in possibilistic decision-making (they have essentially the same form as the possibility distribution but are defined on the strategy set).

Possibilistic expected utility here takes the form of the max-min operation, given that the operation min is the one appropriate to representing the intersection or conjunction (logical multiplication) of fuzzy sets, and the max operation is the one appropriate to representing the set theoretic union or disjunction (the logical sum) of fuzzy sets. In fact possibility and utility are here coincident with fuzzy sets (or fuzzy distributions), the first defined over the set of unforeseen states of the world and the second over the strategy set of each player. Multiplying the possibility of any event by the utility of a decision when that event occurs means, in this context, operating the set-intersection between two fuzzy sets. Summing the expected utility of a decision over different possible events means calculating the set union operation among fuzzy sets. Consider one state of the world at a time, starting from  $\omega_2$  according to the point of

view of player P1. Thus player P1's expected utility from strategy D, given the possibility that in state  $\omega_3$  game G1 or game G2 is played, is as follows

$$U_1(D | \omega_2) = \max\{\min[\pi_{XisG1}(\omega_2), u_1(D|G1)], \min[\pi_{XisG2}(\omega_2), u_1(D|G2)]\} = \\ = \max\{\min(0.2, 0), \min(0.5, 0.21)\} = 0.21$$

where  $u_1(D|G1)$  is player P1's utility from playing the solution strategy D under the hypothesis that the game played is G1,  $\pi_{XisG1}(\omega_2)$  is the possibility of state  $\omega_2$  under the fuzzy information that the game  $G_j$  (i.e. the variable X ranging over the set  $\Omega$  of possible descriptions of interaction situations) played in  $\omega_2$  is G1, and  $U_1(D | \omega_2)$  is player P1's possibilistic expected fuzzy utility from playing D given that  $\omega_2$  has occurred.

In analogous manner we obtain player P1's expected utility from strategy 50 given the possibility that in state  $\omega_2$  game G1 or game G2 is played, which yields  $U_1(50|\omega_2)=0.2$ , and the player P1's expected utility from strategy 50 and strategy D given the possibility that in state  $\omega_1$  game G1 or game G2 is played, which yields  $U_1(50|\omega_1)=0.5$  and  $U_1(D | \omega_1) = 0.21$ , respectively.

Thus in order to maximize his expected utility function contingent on  $\omega_1$ , player P1 has to select the pure strategy 50, but contingently upon  $\omega_2$  P1 must select D.

Hence, we know that if Nature selects state  $\omega_1$  or  $\omega_2$ , then player P1 at the first step of his reasoning process - having at his disposal only the information concerning the possibility that states display certain games - must select 50 or D respectively. This is only his "first glance reasoning", however. It does not account for the fact that if a solution strategy has to be rationally chosen, it must be a best response against the simultaneous choice made by the other party, who faces a similar problem.

## 5.2. Second step in the equilibrium selection process: inferring by default the second player's reasoning and generating an overall conjecture of his choice

Being reflective and rational, even if boundedly so, player P1 endeavours to simulate player P2's reasoning process by introspection. Here default reasoning enters the scene. Player P1 has at his disposal his own model of reasoning at the first step, but there is no evidence that this is also the first step in player's P2 reasoning process. He only knows that in two states of the world  $\omega_1$  and  $\omega_2$  a typical player (himself) has reasoned according to the maximisation rule just mentioned and derived the solution theories to be rationally used in these two cases. No proof exists that any other similarly bounded rational players would not employ similar calculations. Of course such information may be forthcoming later, but as long as the information available to the player is that just described in his model of reasoning, player P1 (at the moment) does not obtain such contrary evidence. Thus, player P1 introduces a default rule of inference by assuming that it is coherent with his knowledge base that player P2 in states like  $\omega_1$  and  $\omega_2$  will perform exactly the same calculations that he performed in his first step in the reasoning process.

Let us define a propositional language by encoding the statements of our theory in the following formulas:

- $\Phi$  := the first-order knowledge base, which encodes statements concerning

- a) the objective description of the states  $\omega_1$  and  $\omega_2$ , including their membership functions to the two fuzzy events  $\underline{G1}$  and  $\underline{G2}$ ;
  - b) the algebraic method to calculate fuzzy expected utility;
  - c) the description of the two solution theory  $\sigma_1$  and  $\sigma_2$ .
- $s :=$  “the strategy D is chosen when the state  $\omega_3$  occurs *and* the strategy 50 is chosen when the state  $\omega_2$  occurs”. This is player P1’s first step scheme of behaviour, and it may be formally written as  $[(\omega_1 \wedge 50) \wedge (\omega_3 \wedge D)]$ .
- Moreover, we encode in our prepositional language the characteristic description of the player  $i$ , for  $i \in \{1,2\}$ , as follows:

- $\Psi(i)$ , whose interpretation is “player  $i$  owns  $\Phi$ ”;
- $\Sigma(i)$ , whose interpretation is “player  $i$  acts according to  $s$ ”;
- $\Gamma(i)$ , whose interpretation is “player  $i$  is rational”.

Let us consider the following set of default clauses: “if player  $i$  owns  $\Phi$  then normally player  $i$  acts according to  $s$ ”, “if player  $i$  is rational then normally he owns  $\Phi$ ”, symbolically written as the set of default  $\Delta = \{\Psi(i) \rightarrow \Sigma(i), \Gamma(i) \rightarrow \Psi(i)\}$ , where  $\Delta$  is the default knowledge base of player P1. Together with the set of formulas constituting the first-order knowledge base of our theory  $\Phi$ , this allows us to define a default theory of the game  $\Theta = \langle \Delta, \Phi \rangle$

In order to infer, as a logical consequence of his default theory  $\Theta = \langle \Delta, \Phi \rangle$ , the decision that will be made by player P2 in each state and his own consequent best response, player P1 needs to combine the expected fuzzy utility calculation with his default knowledge base or conditional knowledge. In particular, player P1 needs to deduce the measure of possibility  $\Pi_2(50|\omega_1)$  on player P2’s choice 50 (i.e. the *plausibility* of the statement “player P2 acts according to  $s$ ” when the state is  $\omega_2$ ) and, similarly, the measure of possibility  $\Pi_2(D|\omega_2)$  on player P2’s choice D. It is in accomplishing this task that the possibilistic logic machinery and its connections with default logic come in.

The basic idea is to encode each default like “if A then normally B”, denoted by  $(A \rightarrow B)$ , into a constraint expressing that the situation where  $(A \wedge B)$  is true has greater plausibility than the one where  $(A \wedge \neg B)$  is true (Benferat et al. 1992, Dubois&Prade 1995a-b, 1996). We follow this approach by encoding the defaults denoted by  $(\Psi(i) \rightarrow \Sigma(i))$  and  $(\Gamma(i) \rightarrow \Psi(i))$  into constraints expressing that the situation where  $(\Psi(i) \wedge \Sigma(i))$  and  $(\Gamma(i) \wedge \Psi(i))$  are true has greater plausibility than the one where  $(\Psi(i) \wedge \neg \Sigma(i))$  and  $(\Gamma(i) \wedge \neg \Psi(i))$  are true. Then we need a qualitative *plausibility* relation  $>_{\Pi}$  for comparing plausibility levels and generating a possibility ranking of situations (Dubois 1986). Following the same line of Dubois&Prade (1995a) and considering the two defaults in player P1’s default knowledge base, the set of defaults  $\Delta = \{\Psi(i) \rightarrow \Sigma(i), \Gamma(i) \rightarrow \Psi(i)\}$  will be represented by the following set of qualitative constraints C:

$$c1: \Psi(i) \wedge \Sigma(i) >_{\Pi} \Psi(i) \wedge \neg \Sigma(i)$$

$$c2: \Gamma(i) \wedge \Psi(i) >_{\Pi} \Gamma(i) \wedge \neg \Psi(i).$$

As numerical counterparts to plausibility relations are possibility measures, the qualitative constraints translate into the following set of ordinal constraints on possibility C’

$$c'1: \Pi(\Psi(i) \wedge \Sigma(i)) > \Pi(\Psi(i) \wedge \neg \Sigma(i))$$

$$c'2: \Pi(\Gamma(i) \wedge \Psi(i)) > \Pi(\Gamma(i) \wedge \neg \Psi(i)).$$

Let  $U$  be the finite set of interpretations of *our propositional language*  $\Psi(i), \Gamma(i), \Sigma(i), \Phi, s, x$ . These interpretations are artificial possible worlds that correspond to formulas formulated in the artificial propositional language just defined, in which the conjunctions  $*\Psi(i)\wedge*\Gamma(i)\wedge*\Sigma(i)$  are true, where  $*$  stands for the presence of the negation sign  $\neg$  or its absence. Hence the models of our formal language are

$$U = \{ u_0: \neg\Psi(i)\wedge\neg\Sigma(i)\wedge\neg\Gamma(i); u_1: \neg\Psi(i)\wedge\neg\Sigma(i)\wedge\Gamma(i); \\ u_2: \neg\Psi(i)\wedge\Sigma(i)\wedge\neg\Gamma(i); u_3: \neg\Psi(i)\wedge\Sigma(i)\wedge\Gamma(i); \\ u_4: \Psi(i)\wedge\neg\Sigma(i)\wedge\neg\Gamma(i); u_5: \Psi(i)\wedge\neg\Sigma(i)\wedge\Gamma(i); \\ u_6: \Psi(i)\wedge\Sigma(i)\wedge\neg\Gamma(i); u_7: \Psi(i)\wedge\Sigma(i)\wedge\Gamma(i) \}.$$

Then the set of ordinal constraints  $C'$  on possibility measures translates into the following set of constraints  $C''$  on the possibility order of models:

$$c''1 : \max(\pi(u_6), \pi(u_7)) > \max(\pi(u_4), \pi(u_5))$$

$$c''2 : \max(\pi(u_5), \pi(u_7)) > \max(\pi(u_1), \pi(u_3)).$$

Now let  $>_\pi$  be a ranking of  $U$ , such that  $u >_\pi u'$  iff  $\pi(u) > \pi(u')$  for each  $u, u' \in U$ . Any finite consistent set of constraints like  $(a \wedge b) >_\Pi (a \wedge \neg b)$  thus induces a partially defined ranking  $>_\pi$  on  $U$  that can be completed according to the principle of minimum specificity. The idea is to try to assign to each world  $u \in U$  the highest possibility level (in forming a well-ordered partition of  $U$ ) without violating the constraints. The ordered partition of  $U$  associated with  $>_\pi$  using the minimum specificity principle can be easily obtained by the following procedure (Dubois&Prade 1995)

$$a. i = 0, F_i = \{\emptyset\}$$

b. As far as  $U$  is not empty, repeat b.1.-b.4.:

$$b.1. i \leftarrow i + 1$$

b.2. Put in  $F_i$  every model which does not appear in the right side of any constraint in  $C''$ ,

b.3. Remove the elements of  $F_i$  from  $U$

b.4. Remove from  $C''$  any constraint containing elements of  $F_i$ .

Let us now apply this algorithm. We obtain:

$$. \quad i=1 : F_1 = \{u_0, u_2, u_6, u_7\}, U = \{u_1, u_3, u_4, u_5\}, C'' = \{\emptyset\}$$

$$. \quad i=2 : F_2 = \{u_1, u_3, u_4, u_5\}, U = \{\emptyset\}, C'' = \{\emptyset\}$$

Finally, the well-ordered partition of  $U$  is:  $\{u_0, u_2, u_6, u_7\} >_\pi \{u_5, u_4, u_1, u_3\}$ .

Note that  $\Sigma(i) \Leftrightarrow u_2 \vee u_7$  and also  $\Sigma(i) \Leftrightarrow u_3 \vee u_6$ . Since  $\Pi(a \vee b) = \max(\Pi(a), \Pi(b))$  and for each  $u \in F_2$  it is true that  $u_2 >_\pi u, u_6 >_\pi u, u_7 >_\pi u$ , then

$$\Pi(\Sigma(i)) \geq \Pi(u \vee u') \text{ for all } u, u' \in U.$$

(1)

In particular also  $\Pi(\Sigma(i)) \geq \Pi(\neg\Sigma(i))$  is satisfied.

Without loss of generality, we can write  $\Pi(\Sigma(i))=\Pi_i(s)$ , with  $i=1,2$ , where  $\Pi_i(s)$  is the measure of possibility that player  $i$  acts according to  $(\omega_1 \wedge 50) \wedge (\omega_2 \wedge D)$ . Moreover, from the axioms governing possibility measures it follows directly that

$$\Pi_i(s)=\Pi_i[(\omega_1 \wedge 50) \wedge (\omega_2 \wedge D)] \leq \min[\Pi_i(\omega_1 \wedge 50), \Pi_i(\omega_2 \wedge D)]$$

and hence directly from (1)

$$\min[\Pi_i(\omega_1 \wedge 50), \Pi_i(\omega_2 \wedge D)] \geq \Pi(-\Sigma(i))= \Pi_i(-s) \quad (2)$$

Consider  $\Pi_i(\omega_1 \wedge 50)$ . From (2) it follows that

$$\Pi_i(\omega_1 \wedge 50) \geq \Pi(-\Sigma(i))= \Pi_i(-s)$$

(3)

Note that, since  $\omega_1$  and  $\omega_2$  are always true, we can write the following relation

$$\neg s \Leftrightarrow \neg((\omega_1 \wedge 50) \wedge (\omega_2 \wedge D)) \Leftrightarrow (\omega_1 \wedge \neg 50) \vee (\omega_1 \wedge \neg D) \vee (\omega_2 \wedge \neg 50) \vee (\omega_2 \wedge \neg D)$$

Therefore

$$\Pi_i(-s) = \max(\Pi_i(\omega_1 \wedge \neg 50), \Pi_i(\omega_1 \wedge \neg D), \Pi_i(\omega_2 \wedge \neg 50), \Pi_i(\omega_2 \wedge \neg D)) \quad (4)$$

From (3) and (4) it follows directly that  $\Pi_i(\omega_2 \wedge 50) \geq \Pi_i(\omega_2 \wedge \neg 50)$ . Since  $\Pi_i(\omega_1) > 0$ , (Benfherat et al. 1997) we have also that

$$N_i(-50 | \omega_1) > 0 \Leftrightarrow \Pi_i(\omega_1 \wedge \neg 50) > \Pi_i(\omega_1 \wedge 50)$$

where  $N_i(-50 | \omega_1)$  is the *conditional necessity measure* which estimates to what extent strategies different from  $s$  have a low degree of possibility of being enacted by player  $i$  given that the state  $\omega_2$  has occurred (Benfherat et al. 1997). Since the second member is false, we conclude that  $N_i(-50 | \omega_1) = 0$ .

By definition  $N_i(-50 | \omega_1) = 1 - \Pi_i(50 | \omega_1)$ , where  $\Pi_i(50 | \omega_1)$  is the *conditional possibility measure* that player  $i$  plays 50 given that the state  $\omega_2$  has occurred; thus we obtain  $\Pi_i(50 | \omega_2) = 1$ . In a similar way we also obtain that  $\Pi_i(D | \omega_2) = 1$ .

It may be concluded that if player P1 uses the default rules represented by the constraints on the possibility ranking of interpretations (artificial worlds) given at the beginning of this section, then the possibility that he assigns to the choices of player P2 reproduces the behaviour that player P1 would have adopted at step 1 in his reasoning. However, in order to produce a reasonable overall guess of players P2's choices, player P1 has to combine two items of vague knowledge, i.e. the first-step fuzzy knowledge about each state being an exemplar of any given possible game, and the second-step default conclusions concerning player P2's first step choice. A reasonable way to do this is by first considering the conjunction of any pair of events like "player P2 chooses strategy 50" and "the game played in the current state is  $G_j$ ," i.e.  $(G_j | \omega_h)$ , for  $j=1,2$  and  $h=1,2$ ; and second by taking the possibility of the resulting conjoint conditioned events, i.e.  $(50 \wedge (G_j | \omega_h))$  - where it happens that when the state  $\omega_h$  is the case, the game played is  $G_j$  and player P2 chooses strategy 50. Given the Bayes-like definition of the possibility of conjoint events in sec.5.1, we may write



$$\Pi_i[50 \wedge (G_j | \omega_h)] = \min[\Pi_i(50 | \omega_h), \pi_{X_{is}G_j}(\omega_h)]$$

where  $i=1,2$ ;  $h=1,2$ ;  $j=1,2$ .

Then we may calculate

$$\Pi_i[50 \wedge (G_1 | \omega_1)] = \min[\Pi_i(50 | \omega_1), \pi_{X_{is}G_1}(\omega_1)] = \min(1, 0.5) = 0.5,$$

as a possibility measure that player  $i$  plays 50 in the game  $G_1$  given the state is  $\omega_1$ , and for the remaining Joint events  $\Pi_i[(D \wedge (G_1 | \omega_2))] = 0.2$ ,  $\Pi_i[(50 \wedge (G_2 | \omega_1))] = 0.4$ ,  $\Pi_i[(D \wedge (G_2 | \omega_2))] = 0.5$ . These possibility measures conclude player P1's attempt to guess the reasoning performed by player P2 at the first step in his reasoning process.

### 5.3. Third step in the selection process: calculating a second hypothetical best response

Given the foregoing guess on player P2 choice, a reasonable way to recalculate expected utility for player P1 is

$$U_1(D | \omega_1) = \max\{\min\{\Pi_2[(50 \wedge G_1) | \omega_1], u_1[D | (G_1 \wedge 50)]\}, \min\{\Pi_2[(50 \wedge G_2) | \omega_1], u_1[D | (G_2 \wedge 50)]\}\} = \max\{\min(0.5, 0), \min(0.4, 1)\} = 0.4$$

where  $u_1[D | (G_1 \wedge 50)]$  is the utility value of player P1 when player P2 plays 50 in the game  $G_1$ . Similarly, it is possible to calculate  $U_1(50 | \omega_1) = 0.5$ ,  $U_1(D | \omega_2) = 0.21$ ,  $U_1(50 | \omega_2) = 0$ .

The expected fuzzy utility of playing 50 when the state  $\omega_1$  occurs is greater than the expected fuzzy utility of playing D when the same state occurs, whereas the expected fuzzy utility of playing D when the state is  $\omega_2$  occurs is greater than that of playing 50 when  $\omega_2$  occurs. Remember that at step two, player P1 stated a set of default rules (the conditional knowledge base  $\Delta$ ) in order - as far as his understanding of the matter was concerned - to account for the reasoning of a "normal" rational player. He must now check whether the conclusions that he reached at step three (by calculating his "second guess" about his highest fuzzy expected utility strategy given his prediction of player P2's choices) are not inconsistent with the default model of "normal" reasoning. Note that a default theory is not monotonic.

To begin with, we can argue that the conditional measure of possibility that player P1 plays 50 given that the state  $\omega_1$  occurs is higher than the conditional measure of possibility that player P1 plays D given that the same state occurs. This is a natural consequence of the reasoning performed by player P1 in his third step. Hence we can say in formulas

$$\Pi_1(50 | \omega_1) > \Pi_1(D | \omega_1).$$

Since  $\Pi_1(50 | \omega_1) = 1 - N_1(\neg 50 | \omega_1)$  and  $\Pi_1(D | \omega_1) = 1 - N_1(\neg D | \omega_1)$ , then  $N_1(\neg D | \omega_1) > N_1(\neg 50 | \omega_1) \geq 0$ .

Moreover, we note that

$$N_1(\neg D | \omega_1) > 0 \Rightarrow \Pi_1(\omega_1 \wedge \neg D) > \Pi_1(\omega_1 \wedge D).$$

Since  $\omega_1 \Leftrightarrow (\omega_1 \wedge \neg D) \vee (\omega_1 \wedge D)$ , then

$$\Pi_1(\omega_1) = \max(\Pi_1(\omega_1 \wedge \neg D), \Pi_1(\omega_1 \wedge D)) = \Pi_1(\omega_1 \wedge \neg D).$$

On the other hand, we have already seen from (3) and (4) that  $\Pi_1(\omega_1 \wedge 50) \geq \Pi_1(\neg s)$ , and that  $\Pi_1(\omega_1 \wedge \neg D)$  is at most as high as  $\Pi_1(\neg s)$ . It follows that

$$\Pi_1(\omega_1 \wedge 50) \geq \Pi_1(\omega_1 \wedge \neg D) = \Pi_1(\omega_1) \quad (5)$$

and by definition

$$\Pi_1(\omega_1 \wedge 50) \leq \min(\Pi_1(\neg 50), \Pi_1(\omega_1)) \leq \Pi_1(\omega_1) \quad (6)$$

Hence, from (5) and (6) we obtain that

$$\Pi_1(\omega_1 \wedge 50) = \Pi_1(\omega_1) \quad (7)$$

Note that, for  $\Pi_1(\omega_1) > 0$ , the conditional possibility measure  $\Pi_1(50|\omega_2)$  is defined (Benfherat et al. 1997) as the greatest solution to the equation  $\Pi_1(\omega_1 \wedge 50) = \min(\Pi_1(50|\omega_1), \Pi_1(\omega_1))$  in accordance with the minimum specificity principle. Therefore

$$\Pi_1(50|\omega_2) = 1 \text{ if } \Pi_1(\omega_2 \wedge 50) = \Pi_1(\omega_2)$$

and

$$\Pi_1(50|\omega_2) = \Pi_1(\omega_2 \wedge 50) \text{ otherwise.}$$

Since from (7) we have  $\Pi_1(\omega_2 \wedge 50) = \Pi_1(\omega_2)$ , thus  $\Pi_1(50|\omega_2) = 1$  and  $N_1(\neg 50|\omega_2) = 1 - \Pi_1(50|\omega_2) = 0$ .

As we saw above, it can be easily checked that for  $\Pi_1(\omega_2) > 0$

$$N_1(\neg 50|\omega_2) > 0 \Leftrightarrow \Pi_1(\omega_2 \wedge \neg 50) > \Pi_1(\omega_2 \wedge 50)$$

Since the first member is false, we conclude that

$$\Pi_1(\omega_1 \wedge 50) \geq \Pi_1(\omega_1 \wedge \neg 50), \quad (8)$$

and

$$\Pi_1(\omega_1 \wedge 50) \geq \Pi_1(\omega_1 \wedge \neg D), \quad (9)$$

and similarly

$$\Pi_1(\omega_2 \wedge D) \geq \Pi_1(\omega_2 \wedge \neg D) \quad (10)$$

and

$$\Pi_1(\omega_2 \wedge D) \geq \Pi_1(\omega_2 \wedge \neg 50). \quad (11)$$

From inequalities (8),(9),(10),(11), comparing the possibility of actions encoded by  $s$  with the possibility of actions encoded by  $\neg s$ , player P1 does not find any inconsistency with the defaults contained in the conditional knowledge base  $\Delta$ . Therefore player P1 has no reason to change his previous conditional knowledge base. Moreover, in this case too, player P1 must continue to use the strategy  $s$  that he calculated as his best choice at the first step in his reasoning process. This means that, for example, insofar as player P2 is predicted by default reasoning to use his strategy  $D$  in the state  $\omega_2$ , player P1's possibilistic best response is strategy  $D$ , because this is player P1's best response calculated given the first-step fuzzy

knowledge on states and the second-step default knowledge on player P2 choices. Of course (D,D) is the equilibrium pair dictated by the solution theory  $\sigma_2$  of game  $G_2$ . However, at this stage the players do not have definite knowledge that the game to be played in the current state is  $G_2$ . Nevertheless, this is the equilibrium pair which tends to emerge from the reasoning process that the players perform in the ongoing game under unforeseen contingencies, when they only know that  $\omega_2$  is the case in point. Conversely, the strategy pair (50, 50) is the emerging solution when they only know that the state  $\omega_1$  is the case in point.

#### 5.4. Further steps: equilibrium selection

Little work is needed to verify that player P2 can be predicted by default to be recalculating his fuzzy expected utility given similar defaults over player P1 reasoning and reaching the symmetrical conclusion that strategy 50 is to be played when  $\omega_1$  is the case and strategy D is to be played when  $\omega_2$  is the case. This would be taken as the fourth step in player P1's default reasoning process. Moreover, if player P1 uses this information in order to recalculate his further best response, it is fairly clear that, given that the data are unchanged, the results cannot change with respect to any states of the world  $\omega_1$  and  $\omega_2$ . That is to say, on applying iterately the same default rules in order to deduce new extensions of the players' theory of the game and new reasonable beliefs based on default inferences, the set of statements of the theory that the players believe no longer changes. Let us state this result as the following

PROPOSITION I: *Each player, by replicating the other player's reasoning process with iterated applications of the same set of defaults belonging to  $\Delta$ , is induced to believe that each of them is going to play the solution strategy-pair  $\sigma_1$  (respectively  $\sigma_2$ ) if the current unforeseen state is  $\omega_1$  (respectively  $\omega_2$ ) even if he does not have incontrovertible proof of the truth of the statement that in the state  $\omega_1$  (respectively  $\omega_2$ ) the game  $G_1$  (respectively  $G_2$ ) is to be played.*

This takes us back to an important feature of an *extension* as defined by Reiter (1980) in his pioneering theory of default reasoning. An extension E is meant as the consistent set of all the formulas derivable from a given first-order knowledge base conjoint to a set of defaults (a conditional knowledge base) - i.e. a default theory  $\Theta = \langle \Delta, \Phi \rangle$  - in such a way that any set of formulas derived by applying default rules to closed formulas belonging to  $\Theta$  or derived from them are already included in E. Let us remind the definition of an extension E as (the minimal) fixed point of the default inference operator  $\Gamma_\Delta$

$$E = \Gamma_\Delta(E)$$

where  $\Gamma_\Delta$  carries sets of closed formulas to new sets of closed formulas by applying the set of defaults  $\Delta$  without introducing any inconsistency into them. This applies to our case because recursive applications of the same defaults to the reasoning of the adversary, in order to generate simulations of the other player's steps in his reasoning process, do not produce new sentences of the theory that contradict the old ones. No changes occur in the calculated best responses, nor changes intervene in the set of beliefs each player holds about the other player's behaviour.

## 5. Conclusions

There is a natural game theoretical interpretation of an extension seen as a fixed point of the default operator applied in order to deduce any internally consistent set of conclusions from the base theory of an agent. Consider that defaults are employed to simulate repeatedly the reasoning of each player by means of the other's models of reasoning, each being concerned with a prediction of the other's behaviour. When the players reach an extension of the knowledge base of the game, which is enlarged by the applications of a basic set of defaults, their beliefs are at a fixed point. Further iterations of the reasoning process, as it is driven by these defaults, will reproduce the same set of beliefs. What emerges is a system of internally *consistent* expectations for each player, where any order layer of a player's beliefs confirms the lower one. A system of reciprocal expectations that are stable and confirm each other at any level, on the other hand, is an appropriate state of players' beliefs in order to say that, contingent on any given state of the world  $\omega \in \Omega$ , they reasonably believe that one equilibrium point will constitute the unique solution of the game. As a consequence, if the hypothesis is accepted that both players reason identically by resorting to the same set of defaults, they will select the same solution concept relative to the given unforeseen state of the world.

Some remarks are in order to point out how stable this result is. *First*, we are assuming inter-subjective vagueness, so that the players  $P_i$  have the same fuzzy-set representations of the relationships between unforeseen states and the *ex ante* defined classes of games  $G_j$ . *Second*, since it is quite obvious that the first-order knowledge base is identical between them, we also assume that they share the same default rules. That is, the model that they employ to guess the other player's reasoning is not only their own model of reasoning attributed by default to the other players but also the same model between them. While the first assumption seems to have some foundation in *fuzzy sets thinking* (Kosko 1993), the second is a simplifying one which should be relaxed in a more general context.

However this need not to be the most serious fragility in the "equilibrium" to which the players converge. The fact that by employing the same default reasoning operator the players eventually reach an *extension* which implies stability of the reciprocal guesses and induces the players to consistently adopt the same solution  $\sigma_j$  does not imply that they are *in fact* playing the proper equilibrium-solution strategy (in the Nash equilibrium sense) of the *effective* game situation that they confronted in the current unforeseen state. By recollecting more information, they could *ex post* discover that they were mistaken in thinking that in state  $\omega_3$ , for example, strategy D was to be used. Thus the solution concept that the players may select as equilibrium of the game under unforeseen contingencies is only a *conjectural* solution concept, in the sense that it is conditional on a given default theory  $\Theta = \langle \Delta, \Phi \rangle$  and it is non-monotonic with respect to the flux of facts that they may learn about the unforeseen states of the world. The possibility of discovering that previous beliefs, apparently justified by the "normal" course of facts, are mistaken is inherent in the non-monotonic nature of default reasoning. Given the players' extension of their theory derived by default rules, they see a solution concept as the rational way to play the game under unforeseen contingencies. They do so, however, without any assurance that *ex post* - when they have accumulated enough experience of the game played under the current state - they will not learn facts showing that in the current state the true game to be played was quite different.

## References

- Bacharach M. 1987. A Theory of Rational decision in Games, *Erkenntnis*. 27:167-190
- Bacharach M. 1994. The Epistemic Structure of a Game. *Theory and Decisions*. 37:7 48.
- Benferhat S.; Dubois D.; and Prade H. 1992. Representing default rules in possibilistic logic. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, 673-684, Cambridge, MA, Oct. 25-29: B. Nebel, C. Rich and W. Swartout editors..
- Benferhat S.; Dubois D.; and Prade H. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, mimeo.
- Binmore K.; and Brandenburger A. 1990. Common Knowledge and Game theory. *Essays on the Foundations of Game Theory*, Basil Blackwell, Oxford, Binmore K. ed.
- Binmore K. 1987/88. Modelling Rational Players, part I and II. *Economics and Philosophy* 3:2-3.
- Brafman R.; and Tennenholtz M. 1997. Modeling Agents as Qualitative Decision-makers. *Artificial Intelligence* 94:217-268.
- Dubois D.; Le Berre D.; Prade H.; and Sabbadin R. 1999. Using possibility theory for modeling qualitative decisions: ATMS-Based algorithms. *Fundamenta Informaticae* 37: 1-39
- Dubois D.; Fargier H.; and Prade H. 1997. Decision-making under ordinal preferences and uncertainty. In *Proceedings of the 13th Conf. on Uncertainty in Artificial Intelligence*, 157-164, Providence, RI, Morgan & Kaufmann, San Francisco, CA.
- Dubois D.; Prade H. 1995a. Possibilistic Logic and Plausible Inference. *Mathematical Models for Handling Partial Knowledge in Artificial intelligence*, Ed. By G.Coletti et Al., Plenum Press, New York.
- Dubois D.; Prade H. 1995b. Possibility Theory as a Basis for Qualitative Decision Theory. In *Proceedings of the 14th Inter. Joint Conf. on Artificial Intelligence (IJCAI'95)*, 1924-1930, Montreal, Canada.
- Dubois D.; and Prade H.; 1996. Fuzzy Sets in Approximate Reasoning: A Personal View, mimeo.
- Dubois D.; and Prade H. 1988. *Possibility Theory*, Plenum Press, New York.
- Dubois D.; Prade H.; Sabbadin R.; 1997. A Possibilistic Machinery for Qualitative Decision. *Qualitative Preference in Deliberation and Practical Reasoning*. Working Notes, March 24-26, Stanford University.
- Fudenberg D; and Levine D. 1998. *Learning in Games*, MIT Press.
- Geffner D. 1996. *Default Reasoning, Causal and Conditional Theories*. Cambridge Mas, MIT Press.
- Ginsberg M. (ed.), (1987) *Readings in Non-monotonic Reasoning*, Morgan Kaufmann Publ. Los Altos.
- Fagin R.; Halpern J.Y.; Moses Y.; Vardi M.Y. 1995. *Reasoning about Knowledge*. Cambridge Mas, MIT Press.
- Harsanyi J.C. 1967-8. Games with Incomplete information Played by *Bayesian* Players, Parts I, II and III. *Management Science* 14:159-82, 320-32, 468-502.
- Harsanyi J.C. 1975. The Tracing Procedure. A Bayesian Approach to Defining a Solution for Non-cooperative Games. *International Journal of Game Theory* 5:61-94.

- Harsanyi J.C. and R.Selten (1988), *A General Theory of Equilibrium Selection*. MIT Press.
- Kreps D. 1990. *Games and Economic Modelling*. Oxford.
- Kosko B. 1993. *Fuzzy Thinking: The New Science of Fuzzy Logic*. Hyperion.
- Lewis D. 1969. *Conventions, A philosophical Study*. Harvard U.P.
- Reiter R. 1980. A Logic for Default Reasoning. *Artificial Intelligence*. 13:81-132.
- Reiter R.; and Cresciuolo G. 1987. On interacting Defaults. *Reading on Non-monotonic Reasoning*. Moragn Kaufmann Publ, Ginsberg ed.
- Sacconi L. 1995. Equilibrio e giustizia, parte II: la selezione del contratto sociale. *Il Giornale degli economisti*. 10/12
- Sacconi L. 2000 *The Social Contract of The Firm*. Springer, Berlin-Heidelberg
- Sacconi L. 2004. Incomplete Contracts and Corporate Ethics: a Game-theoretical Model under Fuzzy Information. LIUC papers n.91, 2001, forthcoming in F.Cafaggi, A. Nicita and U. Pagano (eds.), *Legal Orderings and economic institutions*, Routledge, London (in print).
- Sacconi. L; Moretti S. 2002. Fuzzy Norms, Default Reasoning and Equilibrium Selection in Games under Unforeseen Contingencies, *Serie etica, diritto ed economia*, LIUC Papers n.104.
- Samuelson L. 1997. *Evolutionary Games and Equilibrium Selection*. The Mit Press, Cambridge Mass.
- Sugden R. 1995. A Theory of Focal Points. *Economic Journal*. 105 (430): 533-50.
- Yager R.R. 1979. Possibilistic decision making. *IEEE Trans. on Systems Man and Cybernetics*. 9:388-392.
- Zadeh L.A. 1978. Fuzzy Sets as a Basis for a theory of possibility. *Fuzzy Sets and Systems*. 1:3-28.
- Zimmerman H.J. 1991. *Fuzzy Sets Theory and Its Applications*. 2<sup>nd</sup> revised edition, Kluwer Ac. Press.

Elenco dei papers del Dipartimento di Economia

2000.1 *A two-sector model of the effects of wage compression on unemployment and industry distribution of employment*, by Luigi Bonatti

2000.2 *From Kuwait to Kosovo: What have we learned? Reflections on globalization and peace*, by Roberto Tamborini

- 2000.3 *Metodo e valutazione in economia. Dall'apriorismo a Friedman*, by Matteo Motterlini
- 2000.4 *Under tertiarisation and unemployment.* by Maurizio Pugno
- 2001.1 *Growth and Monetary Rules in a Model with Competitive Labor Markets,* by Luigi Bonatti.
- 2001.2 *Profit Versus Non-Profit Firms in the Service Sector: an Analysis of the Employment and Welfare Implications,* by Luigi Bonatti, Carlo Borzaga and Luigi Mittone.
- 2001.3 *Statistical Economic Approach to Mixed Stock-Flows Dynamic Models in Macroeconomics,* by Bernardo Maggi and Giuseppe Espa.
- 2001.4 *The monetary transmission mechanism in Italy: The credit channel and a missing ring,* by Riccardo Fiorentini and Roberto Tamborini.
- 2001.5 *Vat evasion: an experimental approach,* by Luigi Mittone
- 2001.6 *Decomposability and Modularity of Economic Interactions,* by Luigi Marengo, Corrado Pasquali and Marco Valente.
- 2001.7 *Unbalanced Growth and Women's Homework,* by Maurizio Pugno
- 2002.1 *The Underground Economy and the Underdevelopment Trap,* by Maria Rosaria Carillo and Maurizio Pugno.
- 2002.2 *Interregional Income Redistribution and Convergence in a Model with Perfect Capital Mobility and Unionized Labor Markets,* by Luigi Bonatti.
- 2002.3 *Firms' bankruptcy and turnover in a macroeconomy,* by Marco Bee, Giuseppe Espa and Roberto Tamborini.
- 2002.4 *One "monetary giant" with many "fiscal dwarfs": the efficiency of macroeconomic stabilization policies in the European Monetary Union,* by Roberto Tamborini.
- 2002.5 *The Boom that never was? Latin American Loans in London 1822-1825,* by Giorgio Fodor.
- 2002.6 *L'economia senza banditore di Axel Leijonhufvud: le 'forze oscure del tempo e dell'ignoranza' e la complessità del coordinamento,* by Elisabetta De Antoni.
- 2002.7 *Why is Trade between the European Union and the Transition Economies Vertical?,* by Hubert Gabrisch and Maria Luigia Segnana.
- 2003.1 *The service paradox and endogenous economic growth,* by Maurizio Pugno.
- 2003.2 *Mappe di probabilità di sito archeologico: un passo avanti,* di Giuseppe Espa, Roberto Benedetti, Anna De Meo e Salvatore Espa.

*(Probability maps of archaeological site location: one step beyond*, by Giuseppe Espa, Roberto Benedetti, Anna De Meo and Salvatore Espa).

2003.3 *The Long Swings in Economic Understanding*, by Axel Leijonhufvud.

2003.4 *Dinamica strutturale e occupazione nei servizi*, di Giulia Felice.

2003.5 *The Desirable Organizational Structure for Evolutionary Firms in Static Landscapes*, by Nicolás Garrido.

2003.6 *The Financial Markets and Wealth Effects on Consumption An Experimental Analysis*, by Matteo Ploner.

2003.7 *Essays on Computable Economics, Methodology and the Philosophy of Science*, by Kumaraswamy Velupillai.

2003.8 *Economics and the Complexity Vision: Chimerical Partners or Elysian Adventurers?*, by Kumaraswamy Velupillai.

2003.9 *Contratto d'area cooperativo contro il rischio sistemico di produzione in agricoltura*, di Luciano Pilati e Vasco Boatto.

2003.10 *Il contratto della docenza universitaria. Un problema multi-tasking*, di Roberto Tamborini.

2004.1 *Razionalità e motivazioni affettive: nuove idee dalla neurobiologia e psichiatria per la teoria economica?* di Maurizio Pugno.  
*(Rationality and affective motivations: new ideas from neurobiology and psychiatry for economic theory?* by Maurizio Pugno.

2004.2 *The economic consequences of Mr. G. W. Bush's foreign policy. Can th US afford it?* by Roberto Tamborini

2004.3 *Fighting Poverty as a Worldwide Goal* by Rubens Ricupero

2004.4 *Commodity Prices and Debt Sustainability* by Christopher L. Gilbert and Alexandra Tabova

2004.5 *A Primer on the Tools and Concepts of Computable Economics* by K. Vela Velupillai

2004.6 *The Unreasonable Ineffectiveness of Mathematics in Economics* by Vela K. Velupillai

2004.7 *Hicksian Visions and Vignettes on (Non Linear) Trade Cycle Theories* by Vela K. Velupillai.

2004.8 *Trade, inequality and pro-poor growth: Two perspectives, one message?* by Gabriella Berlofffa and Maria Luigia Segnana.

2004.9 *Worker involvement in entrepreneurial nonprofit organizations. Toward a new assessment of workers? Perceived satisfaction and fairness* by Carlo Borzaga and Ermanno Tortia.



2004.10 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part I): Rational Bargaining and Justification* by Lorenzo Sacconi

2004.11 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part II): Compliance, Reputation and Reciprocity* by Lorenzo Sacconi

2004.12 *A Fuzzy Logic and Default Reasoning Model of Social Norm and Equilibrium Selection in Games under Unforeseen Contingencies* by Lorenzo Sacconi and Stefano Moretti

PUBBLICAZIONE REGISTRATA PRESSO IL TRIBUNALE DI TRENTO