Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Economic Theory 123 (2005) 81–104
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Analogy-based expectation equilibrium

Philippe Jehiel[a,b]

[a] CERAS-ENPC, CNRS (URA 2036), 48 Bd Jourdan 75014 Paris, France
[b] University College, London, UK

Received 7 March 2003; final version received 10 December 2003

Abstract

In complex situations, agents use simplified representations to learn how their environment may react. I assume that agents bundle nodes at which other agents must move into analogy classes, and agents only try to learn the average behavior in every class. Specifically, I propose a new solution concept for multi-stage games with perfect information: at every node players choose best-responses to their analogy-based expectations, and expectations correctly represent the average behavior in every class. The solution concept is shown to differ from existing concepts, and it is applied to a variety of games, in particular the centipede game, and ultimatum/bargaining games. The approach explains in a new way why players may Pass for a large number of periods in the centipede game, and why the responder need not be stuck to his reservation value in ultimatum games. Some possible avenues for endogenizing the analogy grouping are also suggested.

JEL classification: C72; D81

Keywords: Game theory; Bounded rationality; Expectation; Reasoning by analogy

1. Introduction

Standard game theory assumes that players are perfectly rational both in their ability to form *correct* expectations about other players' behavior and in their ability to select *best-responses* to their expectations. The game of chess is a striking example

E-mail address: jehiel@enpc.fr.

0022-0531/$ - see front matter
doi:10.1016/j.jet.2003.12.003

in which the standard approach is inappropriate. In chess, it is clearly impossible to know (learn) what the opponent might do in any event (i.e. at every board position).

In complex situations, players are better viewed as using simplified representations to learn how their environment may react. The main theme of this paper is about understanding the effects of such simplifications onto the limiting outcomes that would emerge after learning has taken place (these will be compactly described through the use of an equilibrium concept). Specifically, I assume the simplifications take the form that players *bundle* nodes at which other players must move into *analogy classes*, and every player only tries to learn the *average* behavior in each analogy class that he considers (as opposed to trying to learn the behavior at every single decision node of his opponents).

The aim of this paper is twofold. The first objective is to propose a solution concept to describe the interaction of players forming their expectations by analogy (as explained above). This will be called the analogy-based expectation equilibrium. The second objective is to analyze the properties of analogy-based expectation equilibria in various strategic interaction contexts, and show how new phenomena may arise. We will also briefly mention some possible avenues for endogenizing the analogy grouping.

The games we consider are multi-stage games with almost perfect information and perfect recall. That is, simultaneous moves and moves by Nature are allowed. But, in any stage, all previous moves are assumed to be known to every player. The partitioning into analogy classes used by the players is given exogenously, and it is viewed as part of the description of the strategic environment.[1] An analogy class α_i of player i is a set of pairs (j,h) such that player j, $j \neq i$, must move at node h. If two elements (j,h) and (j',h') belong to the same analogy class, the action spaces of player j at node h and of player j' at node h' are required to be identically labelled.

Player i's *analogy-based expectation* β_i is player i's expectation about the average behavior of other players in every analogy class α_i considered by player i—we will denote by $\beta_i(\alpha_i)$ the expectation in the analogy class α_i. An *analogy-based expectation equilibrium* is a pair (σ, β) where σ is a strategy profile and β is an analogy-based expectation profile such that two conditions are satisfied. First, for each player i and for each node at which player i must move, player i's strategy σ_i is a *best-response* to his analogy-based expectation β_i.[2] Second, for each player i and analogy class α_i, player i's expectation $\beta_i(\alpha_i)$ *correctly* represents the average behavior in class α_i as induced by the strategy profile σ (where the behavior of player j at node h, $(j,h) \in \alpha_i$, is weighted by the frequency with which (j,h) is visited according to σ—relative to other elements in α_i).[3]

[1] This is similar to the view that information sets are exogenous in the asymmetric information setup of Kreps and Wilson [17]. See however the attempts to endogenize the analogy partitions in Section 6.

[2] More precisely, player i's strategy σ_i is a best-response (at every node where player i must move) to the behavioral strategy that assigns player j to play according to the expectation $\beta_i(\alpha_i)$ at node h, for every (j,h) in the analogy class α_i and for every analogy class α_i.

[3] If no node h such that (j,h) belongs to α_i is ever visited according to σ, (strong) consistency is defined with respect to a small perturbation of σ. (This is in spirit of the definition of sequential equilibrium, see Kreps and Wilson [17].)

I think of consistency as resulting from a learning process in which players would eventually manage to have correct analogy-based expectations (and not as resulting from introspection or calculations on the part of the players). Thus, the approach does not require the players to have any prior knowledge about the analogy classes used by other players (nor about these players' payoffs). It only requires the players to be aware of their own payoffs and of the move structure in the game (i.e., the actions spaces at every decision node of every player).

The approach captures the following aspects of analogy-based reasoning. First, as the common sense of the word 'analogy' suggests, several problems (here expectations) are dealt with together by every player. Second, the correctness of expectations implies that, in any given class, contingencies which are visited more often contribute more to the expectation than contingencies which are visited less often. Accordingly, the behaviors in frequently visited contingencies contaminate the expectation used in all contingencies of the analogy class no matter how often they are visited. The extrapolation (here of the expectation) from more visited to less visited contingencies is—I believe—a key feature of the analogy idea.

In Section 2, I provide a preview of how the solution concept works in a two-player setup in which either normal form game G or G' is played with probability $\frac{1}{2}$, and one player bundles the two games into a single analogy class in order to assess the behavior of his opponent. The example helps clarify how the analogy-based expectation equilibrium *cannot* be viewed as a standard equilibrium (say a Nash–Bayes or sequential equilibrium) of another game, even by varying the information structure (while keeping the payoff and move structures). So the analogy-based expectation equilibrium is an entirely new solution concept whose properties need to be investigated. It is formally defined in Section 3. At the end of Section 3 we make a few preliminary observations. First, we note that an analogy-based expectation equilibrium always exists in finite environments. Second, whenever all players use the finest partitioning as their analogy devices, the strategy profile of an analogy-based expectation equilibrium coincides with a subgame perfect Nash equilibrium. But, otherwise (as illustrated in Section 2) the play of an analogy-based expectation equilibrium need not coincide with that of a subgame perfect Nash equilibrium nor with that of a Nash equilibrium.

In Section 4, we apply the approach to the centipede game [26]. In this game, if a player knows that his opponent will *Take* in the next period he prefers *Taking* in the current period, and in the last period it is a dominant strategy to Take. Thus, in the unique subgame perfect Nash equilibrium, players Take whenever they have to move despite the fact that the payoffs attached to late Take are significantly higher than the payoffs attached to early Take. By contrast, in the analogy-based approach players may keep passing for a large number of periods. More precisely, for a wide range of analogy grouping, there exists an equilibrium in which players only consider Taking whenever the game reaches a late stage.[4]

[4] The phase during which players consider Taking echoes a phenomenon referred to as the end effect in the experimental literature, see Selten and Stoecker [33], but also McKelvey and Palfrey [19] and Nagel and Tang [20].

To illustrate the claim, suppose that each player uses a single analogy class to assess the behavior of his opponent. If players *Pass* most of the time, consistency implies that each player should *expect* his opponent to *Pass* on average with a very large probability. Given such expectations, players find it indeed optimal to *Pass* in all but a few periods toward the end of the game, because they are sufficiently confident that the game will not end at the next node. The key reason why the logic of backward induction fails in this case is that players do not perceive exactly *when* the other player stops *Passing*. Due to their analogy partitioning, players only have a fuzzy perception of their opponent's behavior: they only assess their opponent's average behavior without being able to tell exactly when the opponent Takes. We will comment later on how the analogy-based approach differs from the crazy type approach and other approaches to the finite horizon paradoxes.[5]

In Section 5, we briefly apply the approach to take-it-or-leave-it offer games and to ultimatum and bargaining games. In take-it-or-leave-it situations, we observe that whenever the proposer assesses the probability of acceptance only according to whether his offer is above or below a threshold, the responder may get a payoff that lies strictly above his payoff from refusing the offer (his reservation value), or there may be no agreement despite the presence of positive surplus. In ultimatum and bargaining games, we find that the responder of the ultimatum game may get a significant share of the pie whenever the acceptance decision node of the ultimatum game is bundled with that of the bargaining game for the various alternative offers.

In Section 6, we briefly discuss several possible avenues for endogenizing the analogy partitions. A discussion of the literature and concluding remarks appear in the last section.

2. Preview of the solution concept

Before getting into formal definitions, it may be helpful to illustrate the idea of the solution concept through a simple example.

Two players $i = 1, 2$ must play a normal form game which is either G or G' each with probability $\frac{1}{2}$. Whether the game is G or G', player 1 must choose an action in $\{U, D\}$, and player 2 must choose an action in $\{L, M, R, R'\}$. Players are assumed to know which game G or G' is being played at the time they make their decision. The payoff matrixes of games G and G' are given in the following table:

	L	M	R	R'
U	5, 2	0, 2	2, 4	0, 0
D	4, 3	3, 0	1, 0	2, 0

Game G

	L	M	R	R'
U	3, 0	4, 2	2, 0	1, 1
D	0, 2	5, 2	0, 0	2, 4

Game G'

Both G and G' have a unique Nash equilibrium, which is (U, R) in game G and (D, R') in game G'. Thus, in the standard rationality paradigm, players 1 and 2 would get a payoff of 2 and 4, respectively.

[5] We will also comment on how a similar insight applies to the finitely repeated prisoner's dilemma.

Suppose now that player 1 bundles the two games into a single analogy class in order to assess the behavior of player 2. Player 2 is still assumed to assess the behavior of player 1 separately for games G and G'. The Nash equilibrium is no longer an equilibrium in such an analogy setting. If it were, player 1 would expect player 2 to play R and R' each with probability $\frac{1}{2}$ (remember that the two games are played with equal probability) without being able to distinguish player 2's behavior according to the effective game (G or G') being played. Player 1's best-response to such an expectation would be D in game G $(\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 2>\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 0)$ and U in game G' $(\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 1>\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 2)$, thus being inconsistent with the assumed play of player 1.

The following is an analogy-based expectation equilibrium under the considered profile of analogy grouping:[6]

Strategy profile: Player 1 plays D in game G and U in game G'. Player 2 plays L in game G and M in game G'.

Analogy-based expectations: Player 1 expects player 2 to play L and M each with probability $\frac{1}{2}$ (in his unique analogy class). Player 2 expects player 1 to play D in game G and U in game G'.

To check that the above pair of strategy profile and analogy-based expectations constitute an equilibrium, note that given the strategy profile, players' analogy-based expectations are consistent (G and G' are played with the same probability, thus half of the time player 2 plays L and half of the time he plays M). Given player 1's analogy-based expectation, player 1 chooses D (resp. U) rather than U (resp. D) in game G (resp. G') because $\frac{1}{2}(3+3)>\frac{1}{2}(0+5)$. Given player 1's strategy, player 2's best-response is L in game G and M in game G'.

It is worth noting that player 1 now derives an equilibrium payoff of 3 in game G and 4 in game G'. Thus, player 1 gets strictly more with the coarse analogy grouping than with the fine analogy grouping when player 2 is assumed to be rational.[7] We will make further use of this observation in Section 6 when considering some possible avenues for endogenizing the analogy grouping.

Another important observation is that in the above analogy-based expectation equilibrium, both players 1 and 2 behave differently in games G and G'. Hence (even by varying the payoff matrix specifications), it is not possible to interpret their behavior as resulting from a lack of information as to which game (G or G') is being played. Thus, the analogy-based expectation equilibrium is a new solution concept, and cannot be reduced to an existing solution concept.

Comment: The analogy-based expectation equilibrium cannot either be viewed as a Bayesian equilibrium with *subjective* prior in which player 1 would wrongly believe that player 2 does not distinguish between G and G' (while in fact player 2 would distinguish between the two games). If player 1 were to solve for the game in which player 2 does not distinguish G and G', he could not find it optimal to play D in G and U in G'. If that were the case, player 2 would then strictly prefer playing L

[6] It is the only analogy-based expectation equilibrium in pure strategies.

[7] Of course, such a conclusion could not arise if player 2 had a dominant strategy. Then the coarse analogy of player 1 could only lead him to choose a suboptimal strategy.

(L would give player 2: $\frac{1}{2}3 + \frac{1}{2}0$; M would give: $\frac{1}{2}0 + \frac{1}{2}2$; R would give: $\frac{1}{2}0 + \frac{1}{2}0$; R' would give: $\frac{1}{2}0 + \frac{1}{2}1$), and player 1's best-response to L is U in both G and G'.

3. A general framework

3.1. The class of games

We consider multi-stage games with almost perfect information and perfect recall (see Fudenberg and Tirole [10, Section 3.3.2]). That is, simultaneous moves and moves by Nature are allowed. But, in any stage, all previous moves are assumed to be known to every player.

In the main part of the paper, we will restrict attention to games with a finite number of stages such that, at every stage and for every player (including Nature), the set of pure actions is finite. This class of (finite) multi-stage games with almost perfect information is referred to as Γ.

The standard representation of an extensive form game in class Γ includes the set of players $i = 1, \ldots, n$ denoted by N, the game tree Υ, and the VNM preferences u_i of every player i defined on lotteries over outcomes in the game.

A node in the game tree Υ is denoted by h, and the set of nodes is denoted by H. The set of nodes at which player i must move is denoted by H_i, and for every such node $h \in H_i$, we let $A_i(h)$ denote player i's action space at node h.

3.1.1. Classes of analogy

Each player i forms an expectation about the behavior of other players by pooling together several contingencies in which these other players must move. Each such *pool* of contingencies is referred to as a *class of analogy*. Player i forms an expectation about the *average* behavior in each analogy class that he considers.

Formally, each player i partitions the set $\{(j,h) \in N \times H_j,\ j \neq i\}$ into subsets α_i called analogy classes.[8] The collection of player i's analogy classes α_i is referred to as player i's analogy partition, and it is denoted by An_i. When (j,h) and (j',h') are in the same analogy class α_i, we require that $A_j(h) = A_{j'}(h')$. That is, in two contingencies (j,h) and (j',h') that player i treats by analogy, the action space of the involved player(s) should be the same.[9] The common action space in the analogy class α_i will be denoted by $A(\alpha_i)$. The profile of analogy partitions $(An_i)_{i \in N}$ will be denoted by An.

3.1.2. Strategic environment

A strategic environment in our setup not only specifies the set of players N, the game tree Υ and players' preferences u_i. It also specifies how the various players partition the set of nodes at which other players must move into classes of analogy,

[8] A partition of a set X is a collection of subsets $x_k \subseteq X$ such that $\bigcup_k x_k = X$ and $x_k \cap x_{k'} = \emptyset$ for $k \neq k'$.

[9] More generally, we could allow the players to relabel the original actions of the various players as they wish. From that perspective, $A_j(h)$ should only be required to be in bijection with $A_{j'}(h')$ (as opposed to being equal).

which is summarized in An. A strategic environment is thus formally given by (N, Υ, u_i, An).

3.2. Concepts

3.2.1. Analogy-based expectations

An analogy-based expectation for player i is denoted by β_i. It specifies for every player i's analogy class α_i, a probability measure over the action space $A(\alpha_i)$. This probability measure is denoted by $\beta_i(\alpha_i)$, and $\beta_i(\alpha_i)$ should be interpreted as player i's expectation about the average behavior in class α_i.

3.2.2. Strategy

A behavioral strategy for player i is denoted by σ_i. It is a mapping that assigns to each node $h \in H_i$ at which player i must move a distribution over player i's action space at that node.[10] That is, it specifies for every $h \in H_i$ a distribution—denoted $\sigma_i(h) \in \Delta A_i(h)$—according to which player i selects actions in $A_i(h)$ when at node h. We let σ_{-i} denote the strategy profile of players other than i, and we let σ denote the strategy profile of all players.

3.2.3. Sequential rationality

Given his analogy-based expectation β_i, player i constructs a strategy profile for players other than i that assigns player j to play according to $\beta_i(\alpha_i)$ at node h whenever $(j, h) \in \alpha_i$. (This is the simplest and most natural strategy profile compatible with player i's belief β_i.)[11] The criterion used by player i is that of best-response against this induced strategy profile at every node where player i must move.

More precisely, for every β_i and $j \neq i$, we define the β_i-perceived strategy of player j, $\sigma_j^{\beta_i}$, as

$$\sigma_j^{\beta_i}(h) = \beta_i(\alpha_i) \text{ whenever } (j, h) \in \alpha_i.$$

Given player i's strategy σ_i and given node h, we let $\sigma_i|_h$ denote the continuation strategy of player i induced by σ_i from node h onwards. Similarly, we let $\sigma_{-i}|_h$ and $\sigma|_h$ be the strategy profiles induced by σ_{-i} and σ, respectively, from node h onwards. We also let $u_i^h(\sigma_i|_h, \sigma_{-i}|_h)$ denote the expected payoff obtained by player i when the play has reached node h, and players behave according to the strategy profile σ.

Definition 1 (Criterion). Player i's strategy σ_i is a sequential best-response to the analogy-based expectation β_i if and only if for all strategies σ_i' and all nodes $h \in H_i$;

$$u_i^h(\sigma_i|_h, \sigma_{-i}^{\beta_i}|_h) \geqslant u_i^h(\sigma_i'|_h, \sigma_{-i}^{\beta_i}|_h).$$

[10] Mixed strategies and behavioral strategies are equivalent since we consider games of perfect recall.

[11] In general there may be other strategies that could generate the same analogy-based expectation β_i. Other notions of best-responses to the analogy-based expectations can accordingly be proposed.

3.2.4. Consistency

In equilibrium, I require the analogy-based expectations of the players to be *consistent*. That is, to correspond to the real average behavior in every considered class where the weight given to the various elements of an analogy class must itself be consistent with the real probabilities of visits of these various elements.

I think of the consistency requirement as resulting from a learning process in which players would eventually manage to have correct analogy-based expectations. In line with the literature on learning in games (see Fudenberg and Levine [9]), I distinguish according to whether or not consistency is only required for those analogy classes that are reached with strictly positive probability.[12]

To present formally the consistency idea, let $P^{\sigma}(h)$ denote the probability that node h is reached according to the strategy profile σ.

Definition 2 (Weak consistency). Player i's analogy based expectation β_i is consistent with the strategy profile σ if and only if for all $\alpha_i \in An_i$:

$$\beta_i(\alpha_i) = \left(\sum_{(j,h)\in\alpha_i} P^{\sigma}(h)\cdot\sigma_j(h)\right) \Bigg/ \left(\sum_{(j,h)\in\alpha_i} P^{\sigma}(h)\right), \tag{1}$$

whenever $P^{\sigma}(h) > 0$ for some h and j such that $(j,h)\in\alpha_i$.

The consistency criterion can be interpreted as follows. Suppose that players repeatedly act in the environment as described above. Suppose further that the true pattern of behavior adopted by the players is that described by the strategy profile σ. Consider player i who tries to forecast the average behavior in the analogy class α_i, assumed to be reached with positive probability (according to σ).

The actual behavior in the analogy class α_i is an average of what every player j actually does in each of the nodes h where $(j,h)\in\alpha_i$, that is, $\sigma_j(h)$. The correct weighting of $\sigma_j(h)$ should coincide with the frequency with which (j,h) is visited (according to σ) relative to other elements in α_i. The correct weighting of $\sigma_j(h)$ should thus be $P^{\sigma}(h)/(\sum_{(j,h)\in\alpha_i} P^{\sigma}(h))$, which in turn yields expression (1).

It should be noted that Definition 2 places no restrictions on player i's expectations about those analogy classes that are not reached according to σ. The next definition proposes a stronger notion of consistency (in the spirit of trembling hand or sequential equilibrium, see Kreps and Wilson [17]) that places restrictions also on those expectations.

Formally, let Σ^0 be the set of totally mixed strategy profiles, i.e. strategy profiles σ such that for every player j, for every node $h\in H_j$ at which player j must move, any action a_j in the action space $A_j(h)$ is played with strictly positive probability. For every strategy profile $\sigma\in\Sigma^0$, all analogy classes are reached with positive probability. Thus, there is a unique analogy-based expectation β_i that is *consistent* with σ in the

[12] When it is required for unreached classes, the underlying learning model should involve some form of trembling (or exogenous experimentation). When it is not, trembles are not necessary.

sense of satisfying condition (1) for all analogy classes α_i. Denote this analogy-based expectation by $\beta_i\langle\sigma\rangle$.

Definition 3 (Strong consistency). Player i's analogy-based expectation β_i is strongly consistent with σ if and only if there exists a sequence of totally mixed strategy profiles $(\sigma^k)_{k=1}^{\infty}$ that converges to σ such that the sequence $(\beta_i\langle\sigma^k\rangle)_{k=1}^{\infty}$ converges to β_i.

3.2.5. Solution concept

In equilibrium, I require that at every node players play best-responses to their analogy-based expectations (sequential rationality) and that expectations are consistent. I define two solution concepts according to whether or not consistency is imposed on analogy classes that are not reached along the played path. And I refer to a pair (σ, β) of strategy profile and analogy-based expectation profile as an *assessment*.

Definition 4. An assessment (σ, β) is an analogy-based expectation equilibrium (resp. a self-confirming analogy-based expectation equilibrium) if and only if for every player $i \in N$,

1. σ_i is a *sequential best-response* to β_i and
2. β_i is *strongly consistent* (*resp. consistent*) with σ.

It may be worth stressing a few notable differences between an analogy-based expectation equilibrium and a sequential equilibrium of an extensive form game with incomplete information. First, observe that an analogy partition of, say player i, is a partition of the nodes where players *other than i* must move. It is thus of a different nature than player i's information structure in a game with incomplete information which refers to a partition of the nodes where player *i himself* must move.[13] Second, observe the different nature of player i's analogy-based expectation $\beta_i(\cdot)$ and of player i's belief system in extensive form games with incomplete information. Here $\beta_i(\alpha_i)$ is an expectation (or belief) about the *average behavior* of players other than i in class α_i. It is *not* a belief, say, about the likelihood of the various elements (j, h) pooled in α_i. Finally, note that in the analogy setup, the same *expectation* is used to assess the behavior of the opponent(s) in two elements of an analogy class. By contrast, in the incomplete information setup a player *behaves* in the same way at two nodes of a common information set. These differences can be seen in the simple setup developed in Section 2, thus revealing that the analogy-based expectation equilibrium is an entirely new solution concept that cannot be interpreted as a

[13] In the incomplete information setup, player i's action spaces are required to be the same at two nodes of a common information set. Compare with our requirement that $A_j(h) = A_{j'}(h')$ whenever (j, h) and (j', h') are in the same analogy class.

sequential equilibrium of a modified game (possibly with a different information structure).

3.3. Preliminary results

Two simple observations follow. The first one shows the relation of the analogy-based expectation equilibrium to the subgame perfect Nash equilibrium when all players use the finest partitioning as their analogy device. The second one shows the existence of analogy-based expectation equilibria in finite environments.

Proposition 1. *Consider an environment (N, Υ, u_i, An) in which all players use the finest analogy partitioning.*[14] *Then if (σ, β) is an analogy-based expectation equilibrium of (N, Υ, u_i, An), σ is a subgame perfect Nash equilibrium of (N, Υ, u_i).*

Proof. When players use the finest analogy partitioning, strong consistency of β_i with respect to σ implies that $\sigma_{-i}^{\beta_i} = \sigma_{-i}$. Proposition 1 then follows from Definition 1. $\square$

When at least one player, say player i, does not use the finest partition as his analogy device, the play of an analogy-based expectation equilibrium need not correspond to that of a subgame perfect Nash equilibrium. This is because in an analogy-based expectation equilibrium (σ, β), player i's strategy σ_i is required to be a best-response to $\sigma_{-i}^{\beta_i}$. But, $\sigma_{-i}^{\beta_i}$ need not (in general) coincide with σ_{-i} as in a subgame perfect Nash equilibrium. This has been illustrated in Section 2 and will be further illustrated throughout the paper.

Proposition 2 (Existence). *Every finite environment (N, Υ, u_i, An) has at least one analogy-based expectation equilibrium.*

Proof. The strategy of proof is the same as that for the existence proof of sequential equilibria [17]. I mention the argument, but for space reasons I do not give the details of it. First, assume that in every node $h \in H_i$, player i must choose every action $a_i \in A_i(h)$ with probability no smaller than ε (this is in spirit of Selten [32]).[15] It is clear than an analogy-based expectation equilibrium with such additional constraints must exist. Call $(\sigma^\varepsilon, \beta^\varepsilon)$ one such profile of strategies and analogy-based expectations. By compactness properties (which hold in the finite environment case), some subsequence must be converging to say (σ, β), which is an analogy-based expectation equilibrium. $\square$

[14] We say that all players use the *finest* analogy partitioning if there are no i, (j,h), $(j',h') \neq (j,h)$ and $\alpha_i \in An_i$ such that $(j,h) \in \alpha_i$ *and* $(j',h') \in \alpha_i$.

[15] This requires amending Definition 1 to incorporate such constraints in the maximization programs.

4. Finite horizon paradox

Consider the centipede game *CP* depicted in Fig. 1 (see Rosenthal [26]). Two players $i = 1, 2$ move in alternate order starting with player 2. There are $2K$, $K \geqslant 2$, decision nodes as labelled in the figure. At each node, the player whose turn it is to move, say player i, may either *Take* in which case this is the end or *Pass*, i.e. $A_i = \{Pass, Take\}$. The game also ends when player 1 *Passes* at node $N_1^{(1)}$. The scalars a_t and b_t define the payoffs at each terminal node. These scalars are assumed to be non-negative and satisfy (whenever applicable):

$$a_{2k-1} > a_{2k+1} > a_{2k},$$
$$b_{2k-2} > b_{2k} > b_{2k-1}. \tag{2}$$

These conditions guarantee that (1) the unique subgame perfect Nash equilibrium (SPNE) of *CP* is such that players Take whenever they have to move (this follows from $a_{2k+1} > a_{2k}$ and $b_{2k} > b_{2k-1}$), and (2) in any period $t \leqslant 2K - 2$, the player whose turn it is to move is better off if *Take* occurs two periods later, i.e. in period $t + 2$, than if it occurs in the current period t (this follows from $a_{2k-1} > a_{2k+1}$ and $b_{2k-2} > b_{2k}$). Note that we do not impose any lower bound on a_{2k} or b_{2k-1}. Thus, the set of conditions is compatible both with the original Rosenthal's version in which both players' payoffs increase as the game moves further and with the dollar game version (see Reny [25]) in which the player who does not Take consistently gets the same low payoff, say 0, throughout the game. It is also possible that the payoff obtained by the player who does not Take decreases as the game moves further—we refer to the latter version as the decreasing loser's payoff version. Some of our discussion will differentiate between these various payoff specifications.

4.1. The coarsest analogy grouping

Our main insight is that for a wide range of analogy grouping, players may keep *Passing* for a large number of periods in equilibrium. To illustrate the claim, we first consider the situation in which both players use the coarsest analogy partition. That is, each player i pools together *all* the nodes $N_j^{(k)}$ at which player j, $j \neq i$ must move

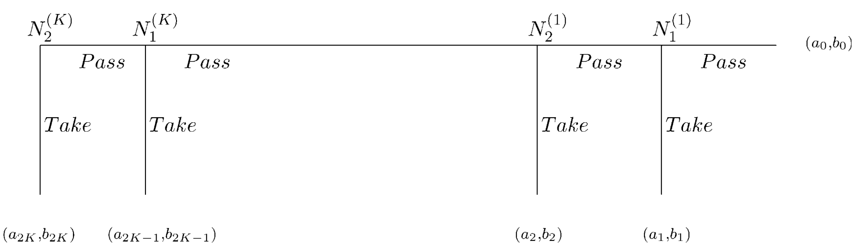

Fig. 1. The centipede game.

into a single class of analogy α_i:

$$\alpha_i = \{(j, N_j^{(k)}), 1 \leqslant k \leqslant K\}.$$

In the environment with coarsest analogy grouping, we have:

Proposition 3. *Suppose that for all* $k \geqslant 1$,

$$\tfrac{1}{2} a_{2k-1} + \tfrac{1}{2} a_{2k} > a_{2k+1} \quad \tfrac{1}{2} b_{2k-2} + \tfrac{1}{2} b_{2k-1} > b_{2k}. \tag{3}$$

There is a unique self-confirming analogy-based expectation equilibrium in which Take never occurs in the first two periods: It is such that players keep Passing throughout the game except in the last node $N_1^{(1)}$ *at which player* 1 *Takes.*

Proposition 3 makes two points. First, there is an equilibrium outcome in which Take occurs only at the final decision node $N_1^{(1)}$. Second, there is no equilibrium (whether in pure or in mixed strategy) in which Take occurs in the middle of the game (i.e. between nodes $N_1^{(1)}$ and $N_1^{(K)}$). Thus either Take occurs immediately (the SPNE is an equilibrium in this setting too[16]) or it occurs at the very end of the game.[17]

Proof of Proposition 3. First, it is readily verified that Take at the last node $N_1^{(1)}$ can be sustained as an equilibrium. If players so behave, weak consistency implies that (1) player 2 should expect player 1 to Pass on average with probability $\frac{K-1}{K}$, i.e. $\beta_2(\alpha_2) = \frac{K-1}{K} \cdot Pass + \frac{1}{K} \cdot Take$ (player 1 has K decisions nodes $N_1^{(k)}$ each met with probability 1, i.e. $P^\sigma(N_1^{(k)}) = 1$, and player 1 Takes only in one of these), and (2) player 1 should expect player 2 to Pass (on average) with probability 1, i.e. $\beta_1(\alpha_1) = Pass$ (since player 2 passes throughout the game). Clearly, player 1 is playing a best-response to his expectation as he gets his maximal payoff. By Passing always, player 2 is also playing a best-response to his analogy-based expectation. Indeed given that $\beta_2(\alpha_2) = \lambda \cdot Pass + (1-\lambda) \cdot Take$ with $\lambda > \frac{1}{2}$ and $\frac{1}{2} b_{2k-2} + \frac{1}{2} b_{2k-1} > b_{2k}$, at every node $N_2^{(k)}$ player 2 strictly prefer Passing (and Taking at $N_2^{(k-1)}$ if $k > 1$) to Taking in the current node.

Clearly, if the game reaches the last decision node $N_1^{(1)}$, player 1 must Take (this is his best strategy whatever his expectation). Thus, Take must occur at some decision node. But, it is easy to see that there is no equilibrium in *pure* strategies in which Take can occur in the course of the game, say at node $N_i^{(k)}$ where $N_i^{(k)} \neq N_2^{(K)}$ and $N_1^{(1)}$. The reason is that weak consistency would then imply that player i should expect player j to Pass with probability 1 (since he would never Take on the equilibrium path), and player i's best-response to such an expectation is definitely

[16] The corresponding analogy-based expectations are that players Take with probability 1.

[17] There is also an equilibrium in mixed strategy in which each player i, $i = 1, 2$ plays in mixed strategies at his first decision node $N_i^{(K)}$, and Takes with probability 1 thereafter.

not to Take at node $N_i^{(k)}$, since player i could increase his payoff by Passing one more time, say.

Proposition 3 also rules out the possibility of a mixed strategy self-confirming analogy equilibrium in which each player i Passes with probability 1 at his first decision node. Again, if players Pass with probability 1 at their first decision node, weak consistency implies that the analogy-based expectation equilibrium of player i should satisfy

$$\beta_i(\alpha_i) = \lambda^i \cdot Pass + (1 - \lambda^i) \cdot Take \quad \text{with } \lambda^i \geqslant \tfrac{1}{2}.$$

(This is because the lowest Pass rate is obtained when player i Takes with probability 1 at his second decision node $N_i^{(2)}$ and the corresponding Pass rate is $\frac{1}{2}$.) Given our assumption on payoffs, player i's best-response to β_i is either to Pass always for player 2 or to Pass till node $N_1^{(1)}$ for player 1. □

4.1.1. Length effect and coarse partitioning

Whenever condition (3) is not met, Proposition 3 does not apply, but we still get similar insights for long enough versions of the game. More precisely, assume that for all $k \geqslant 1$, $a_{2k-1} > \rho \cdot a_{2k+1}$ and $b_{2k-2} > \rho \cdot b_{2k}$ with $2 > \rho > 1$ and let m be such that $\rho \cdot \frac{m}{m+1} > 1$. (Observe that m is set independently of K.)

It is readily verified that if both players Pass in their first m decision nodes then in any equilibrium of the coarsest analogy grouping environment players must keep Passing till the last node $N_1^{(1)}$. Thus, for long enough versions of the game, either Take occurs in the initial phase of the game (in the first $2m$ decision nodes) or otherwise the game must proceed to the very last decision node of the game at which node player 1 Takes.

It is instructive to illustrate how increasing the length of the game may facilitate the possibility of Passing toward the end of the game. Consider the following dollar game version ($a_{2k-2} = b_{2k-1} = 0$) in which for all $k \geqslant 1$, $a_{2k-1} = \rho \cdot a_{2k+1}$ and $b_{2k-2} = \rho \cdot b_{2k}$ where $\rho = 1.75$ and $a_1 = b_0 = 100$. On can check that:

- When $K = 2$, the highest Pass rate equilibrium is such that each Player i Passes with probability 1 at node $N_i^{(2)}$, player 2 Passes with probability $\frac{3}{4} = \rho - 1$ at node $N_2^{(1)}$ and player 1 Takes at node $N_2^{(1)}$.[18]
- When $K \geqslant 3$, the above behavioral strategies at nodes $N_i^{(k)}$, $k = 1, 2$ cannot be part of an equilibrium and the highest Pass rate equilibrium is such that players Pass till the last node $N_1^{(1)}$ at which node player 1 Takes.[19]

[18] The corresponding belief of player 2 is that player 1 Passes with probability $\frac{1}{1+3/4} = \frac{4}{7}$ which multiplied by ρ is exactly 1. Thus, player 2 is indeed indifferent between Passing and Taking at node $N_1^{(1)}$. The belief of player 1 is that player 2 Passes with probability $\frac{1+3/4}{2} = \frac{7}{8}$ which multiplied by ρ is strictly larger than 1. Thus, player 1 strictly prefers Passing at node $N_1^{(2)}$ (and Taking at node $N_1^{(1)}$).

[19] When the outcome is Take at the last node, player 2's belief is that player 1 Passes with probability $\frac{K-1}{K} \geqslant \frac{2}{3}$ whenever $K \geqslant 3$. Since $\frac{2}{3} \cdot \rho > 1$ the best response for player 2 is to Pass always.

It is interesting to note that the patterns of behavior in the subgame starting at node $N_2^{(2)}$ are not the same according to whether $K = 2$ or $K \geqslant 3$, even though the subgames are identical in terms of move and payoff structure. The reason is that for longer versions early Pass behaviors (at nodes $N_i^{(k)}$, $k \geqslant 3$) typically raise the expectation of the average Pass rate all over the game, and thus longer versions of the game in which players keep Passing till node $N_2^{(2)}$ result in higher Pass rate expectations, which in turn induces more Pass behavior in the subgame starting at $N_2^{(2)}$. This feature will be contrasted with the insight arising with the crazy type approach.

4.2. Arbitrary analogy grouping

So far we have only considered the coarsest analogy partitions. More generally:

Proposition 4. *Suppose that condition* (3) *holds. Consider an arbitrary profile of analogy partitions* (An_1, An_2), *and let* $N_i^{(k)}$ *be the latest node*[20] *such that the analogy class to which* $N_i^{(k)}$ *belongs is not a singleton.*[21] *Take at node* $N_i^{(k)}$ *can be sustained as an equilibrium, and it is the equilibrium with highest Pass rate.*

Proof. We rule out the case where both players use the finest partition for which Proposition 4 trivially applies. We let α_j be the non-singleton class that contains $N_i^{(k)}$. Since all nodes following $N_i^{(k)}$ are in a singleton class, backward induction arguments imply that players Take whenever they have to move after node $N_i^{(k)}$. At node $N_i^{(k)}$ player i also finds it optimal to Take given that his expectation must be that player j will Take in the next node. Thus, the highest possible Pass rate is obtained when Take occurs at node $N_i^{(k)}$. Finally, if Take occurs at node $N_i^{(k)}$ consistency implies that

$$\beta_j(\alpha_j) = \lambda^j \cdot Pass + (1 - \lambda^j) \cdot Take \quad \text{with } \lambda^j \geqslant \tfrac{1}{2}.$$

(This is because node $N_i^{(k)}$ is bundled with at least one earlier node $N_i^{(k')}$, $k' > k$ and at this node player i Passes with probability 1; also each node up to $N_i^{(k)}$ is met with probability 1.) Under condition (3), such an expectation guarantees that player j finds it optimal to Pass at the node preceding $N_i^{(k)}$ and it is readily verified that Passing till node $N_i^{(k)}$ and Taking from node $N_i^{(k)}$ onwards constitutes a best-response to the corresponding analogy-based expectations. $\square$

[20] That is, consider such a node $N_i^{(k)}$ (belonging to a non-singleton α_j) with smallest k and choose player 1 in case of tie between the two players.

[21] If there is no such node, let $N_i^{(k)} = N_2^{(K)}$.

Thus, for arbitrary analogy grouping, to the extent that some node close to the end is not in a singleton analogy class, there always exists an equilibrium with a very high Pass rate.

Of course, there is still the possibility that Take occurs at an early stage for some equilibria. But, if we impose the extra requirement that all analogy classes must be reached in equilibrium,[22] we get:

Proposition 5. *Suppose that condition* (3) *holds. Consider a profile of analogy partitions and let* (σ, β) *be a corresponding analogy-based expectation equilibrium in pure strategies. If* all *analogy classes of both players are reached with positive probability according to* σ, *then the equilibrium outcome is that player* 1 Takes *in the last node* $N_1^{(1)}$.

Proof. Take at node $N_1^{(1)}$ is a possible equilibrium outcome when players use the coarsest partition (see Proposition 3). Since all classes of both players are then reached with positive probability, this outcome can be sustained in the way required by the proposition.

Suppose that another outcome, i.e. player i Takes at node $N_i^{(k)}$ with $(i,k) \neq (1,1)$, were to emerge with the same requirements.

First, it cannot be that this outcome corresponds to the subgame perfect Nash equilibrium outcome, since then no node $N_1^{(k)}$ would be reached, and thus at least one of the analogy classes of player 2 would not be reached in equilibrium.

If player i were to Pass at node $N_i^{(k)}$ this would lead to node $N_j^{(k')}, j \neq i$, with $k' = k$ if $i = 1$ and $k' = k - 1$ if $i = 2$. Since node $N_j^{(k')}$ is not reached in equilibrium and since all analogy classes must be reached with positive probability, it must be that there is an analogy class α_i of player i such that $(j, N_j^{(k')}) \in \alpha_i$ and $(j, N_j^{(k'')}) \in \alpha_i$ where $k'' < k'$ (nodes $N_j^{(k'')}$ with $k'' > k'$ are not reached).[23] Since at any node $N_j^{(k'')}$ with $k'' < k'$ player j Passes with probability 1 (remember that Take at node $N_i^{(k)}$ is the assumed outcome), it must be that the analogy-based expectation of player i satisfies

$$\beta_i(\alpha_i) = \lambda^i \cdot Pass + (1 - \lambda^i) \cdot Take \quad \text{with } \lambda^i \geqslant \tfrac{1}{2}.$$

But given this expectation (and given condition (3)), Taking at node $N_i^{(k)}$ cannot be a best-response to β_i (at node $N_i^{(k)}$, player i should strictly prefer Passing rather than Taking). This leads to a contradiction. □

Observe that Proposition 5 does not pin down a single analogy grouping. Yet, it guarantees that there must be a very high Pass rate on the equilibrium path of any pure strategy equilibrium such that all classes are reached with positive probability.

[22] Such a requirement may be related to the psychological trait that all contingencies should eventually be related to some contingencies that have some familiarity to the player.

[23] There exists at least one such node because $(i,k) \neq (1,1)$.

4.3. Link to other approaches

We now review how our approach differs from the existing approaches to the finite horizon paradoxes:

Crazy type: The so called crazy type approach [16] is definitely the most influential and popular one. Applied to the centipede game, it can be described as follows:[24] Suppose that with small but positive probability players behave mechanically and Pass in all periods. The key insight of the crazy type approach is that players even when rational will Pass most of the time. The reason is that if a rational player were to Take at the start say, he could easily convince (according to the sequential equilibrium) his opponent that he is of the type who Passes always by Passing one more time, and such a change of belief would, of course, be very beneficial to the player.

It should be noted that the rational agents of the crazy type approach are highly sophisticated in their ability to make inferences from observed behaviors onto the likelihood of their opponents' types. It is thus not a model of bounded rationality in contrast to the current approach (which assumes that players do not exactly identify when their opponent stops Passing).

There are a few notable behavioral differences between the two approaches. While in the analogy-based approach, it is never possible that player 1 Passes at the last decision node $N_1^{(1)}$, the crazy type approach does allow for this possibility when player 1 is truly the type who Passes always (and this is quite essential for the argument to work). Furthermore, increasing the number of periods does not have the same effect in the two approaches. In the analogy approach we have seen that longer versions of the centipede game may increase the Pass rate in the last few nodes (see above). By contrast, in the crazy type approach, if the rational types are assumed to Pass till node $N_2^{(2)}$ the play in the subgame starting at that node should be independent of the total duration of the game (because the equilibrium belief as to whether or not one faces a crazy type is not affected by this total duration).

Complexity: Another approach is that players are constrained in the complexity of the strategies they can employ Neyman [21].[25] To illustrate the approach, assume that player 1 is constrained to use a strategy describable by a machine with no more than J states, $J < K$. A machine with at most J states does not allow player 1 to differentiate his behavior in periods t, $t \leqslant 2(K-J)-1$. Thus, whenever player 1' strategy is constrained to use no more that J states, and $a_{2k-1} \leqslant a_0$ for all k, $K \leqslant k < K-J$, Passing for both players in all periods is an equilibrium. Note that the

[24] It was originally applied to the finitely repeated prisoner's dilemma. To apply our approach to the finitely repeated prisoner's dilemma, suppose, for example, that players use two analogy classes categorizing histories according to whether or not some opportunistic behavior was played earlier in the interaction, Cooperating in all but the last period is an analogy-based expectation equilibrium (see Jehiel [14] for details).

[25] See also Rubinstein [27], Abreu and Rubinstein [1] and Rubinstein [30, Chapter 9] for a related approach in which players not only care about their material payoff, but also about the complexity of their strategy.

prediction in this case is that both Players Pass with probability 1 at each of their decision nodes (which can never arise in the analogy-based approach in which Take must occur at least at node $N_1^{(1)}$). Note also that in the dollar game version of the game we have that $a_0 = 0$, and $a_{2K-1} > a_0$. Thus, Passing in all periods is not an equilibrium in the bounded complexity approach. More generally, in the dollar game or decreasing loser's payoff specifications, the only possible equilibrium outcome of the bounded complexity approach is that player 2 Takes at his first decision node. This should again be contrasted with the analogy-based approach (see Proposition 3 which applies equally to all such payoff specifications).

Imperfect recall: A third approach is that players have imperfect recall and do not remember exactly at which stage of the game they currently are (see Piccione and Rubinstein [23] and Dulleck and Oechssler [5] for an application to the centipede game). For the sake of illustration, assume that each player $i = 1, 2$ has a unique memory set consisting of all his decisions nodes. That is, player i has no idea at which node $N_i^{(k)}$, $k = 1, \ldots, K$ he currently is. For K large enough, an equilibrium in this setting is that each player i Passes with probability 1 in his unique memory/information set.[26] Note that the prediction is that both players Pass with probability 1 in all periods. This prediction arises because players do not perceive that there is an end, which in turn forces the players to have the same behavior independently of the decision node. This is in sharp contrast with the analogy-based approach in which players do perceive that there is an end no matter how they categorize their opponent's nodes into analogy classes (for example, in the coarsest partitioning player 1 always Takes at his last decision node $N_1^{(1)}$, thus revealing that he does perceive that there is an end).

5. Application to bargaining and ultimatum games

In this section, we briefly apply the analogy-based approach to take-it-or-leave-it and bargaining games.

5.1. Take-it-or-leave-it

Consider the following situation. There are two players $i = 1, 2$ and a pie of size 1. Player 1 makes a partition offer $(x, 1 - x)$, $x \in [0, 1]$ to player 2 who may either accept or reject it.[27] If he accepts, players 1 and 2 get x and $1 - x$, respectively. If player 2 rejects the offer, player 1 gets 0 and player 2 gets an outside option payoff equal to v^{out}, where $0 < v^{\text{out}} < 1$.

[26] The point is that player i cannot adjust the best time for Taking, as he does not know at which $N_i^{(k)}$ he currently is. He prefers Passing always in this case.

[27] The action space of player 1 in this example is continuous (which is not covered by the framework of Section 3). The analysis presented below can be viewed as corresponding to the limit of the finite grid case as the grid becomes finer and finer.

Standard analysis suggests that player 1 will propose $(1 - v^{\text{out}}, v^{\text{out}})$ and that player 2 will accept it. When player 1 forms his expectation about player 2's probability of acceptance by analogy, we now show that it may well be that either player 1 makes a much more generous offer than v^{out} to player 2 or that player 1 makes an offer that is rejected by player 2 depending on the partitioning.

Specifically, a node at which player 2 must move can be identified with x where $(x, 1-x)$ is the offer made by player 1. We assume that player 1 partitions the set of $(2, x)$ into two classes:[28]

$$\alpha_1^{\text{low}} = \{(2,x) | \bar{x} < x \leqslant 1\},$$
$$\alpha_1^{\text{high}} = \{(2,x) | 0 \leqslant x \leqslant \bar{x}\},$$

where α_1^{low} (resp. α_1^{high}) corresponds to the class of outrageous (resp. generous) offers.

Proposition 6. (1) *When* $1 - \bar{x} < v^{\text{out}}$, *any analogy-based expectation equilibrium is such that there is no agreement*: *player* 1 *gets* 0, *player* 2 *gets* v^{out}. (2) *When* $v^{\text{out}} < 1 - \bar{x}$, *there is a unique analogy-based expectation equilibrium*: *player* 1 *proposes* $(\bar{x}, 1 - \bar{x})$, *and player* 2 *accepts*.

Proof. The analogy-based expectation of player 1 is of the form $\beta_1(\alpha_1^r) = \lambda^r \cdot$ 'Accepts' $+(1-\lambda^r)\cdot$ 'Rejects' with $r = \text{low, high}$. If $\lambda^{\text{high}} > 0$ (resp. $\lambda^{\text{low}} > 0$), player 1's best-response to β_1 cannot be to offer $(x, 1-x)$ with $x < \bar{x}$ (resp. $\bar{x} < x < 1$). (1) When $1 - \bar{x} < v^{\text{out}}$, neither $(1, 0)$ nor $(\bar{x}, 1 - \bar{x})$ are acceptable by player 2. Only a disagreement can occur. (2) When $v^{\text{out}} < 1 - \bar{x}$, $\lambda^{\text{high}} = 1$, $\lambda^{\text{low}} = 0$, player 1 proposing $(\bar{x}, 1 - \bar{x})$ and player 2 accepting any offer $(x, 1-x)$ with $1 - x \geqslant v^{\text{out}}$ gives rise to an analogy-based expectation equilibrium. It is also easy to see that there is no other analogy-based expectation equilibrium. (For example, a disagreement cannot be part of an equilibrium, because strong consistency would force $\lambda^{\text{high}} = 1$. Thus, offering $(\bar{x}, 1 - \bar{x})$ is a better option for player 1 than just opting out, leading to a contradiction.) $\square$

The intuition underlying this result is rather immediate. The analogy-based approach leads the proposer to have the same expectation about the acceptance rate for all offers lying in a given analogy class. Thus, conditional on making an offer in a given analogy class, best-response will induce the proposer to pick the least generous offer among these. So analogy grouping has the effect of (endogenously) discretizing the action space of the proposer (to the lower extreme points of his analogy classes), which in turn explains why there may be disagreement or the responder need not be stuck to his reservation utility.[29]

[28] The intervals are closed as indicated to guarantee the existence of an equilibrium.

[29] The approach should be contrasted with an approach (in the spirit of Dow [4] or Rubinstein [29]) in which the responder would have a coarse perception of the proposer's offers. In such an approach, the responder has the same behavior in terms of acceptance for all offers lying in a given information/perception set because he does not distinguish them. By contrast, in the analogy-based approach the decision of acceptance need not be the same for two offers corresponding to the same analogy class.

5.2. Bargaining vs ultimatum

The situation is now as follows. There are two players $i = 1, 2$ and a pie of size 1. The game governing the division of the pie is either an alternating offer bargaining game à la Rubinstein R (with a discount factor assumed to be close to 1 for both players), or an ultimatum game U_1 in which player 1 makes a take-it-or-leave-it offer to party 2 (who is assumed to get 0 if he refuses the deal) or an ultimatum game U_2 symmetrically defined in which player 2 makes the offer. Each game arises with probability $\frac{1}{3}$, and to simplify the exposition, we assume that the offers are constrained to be $(1/4, 3/4)$, $(1/2, 1/2)$ or $(3/4, 1/4)$.

Standard analysis would tell that in U_i the proposer should propose $\frac{3}{4}$ for himself, which should be accepted by the responder. Regarding bargaining game R, an equilibrium is that players always propose the equal split $(\frac{1}{2}, \frac{1}{2})$, and it is accepted.

Suppose now that in order to assess the acceptance probability players do not distinguish between ultimatum and bargaining games whereas (unlike in Section 5.1) they do distinguish between the acceptance probability attached to the various possible offers.

It is readily verified that the path induced by the above SPNE can no longer arise with such an analogy grouping. If it were, consistency would imply that each player i should expect that an offer of $3/4$ for himself should be accepted (in ultimatum *and* bargaining games)—such an offer would only be made in ultimatum games where it would be accepted. But, with such an expectation a player would definitely not propose an equal split offer in the bargaining game R (he would rather propose $3/4$ for himself).

A possible analogy-based expectation equilibrium in this setting is that (1) players always propose an equal split $(\frac{1}{2}, \frac{1}{2})$ in both the bargaining game and the ultimatum game; (2) all offers are accepted in the ultimatum game; the equal split $(\frac{1}{2}, \frac{1}{2})$ offer and any more generous offer is accepted in the bargaining game; (3) when a player proposes $3/4$ for himself he expects the offer to be rejected with probability $1/2$,[30] and when a player proposes $1/2$ or less for himself he expects the offer to be accepted (with probability 1).

Observe that this setting too explains why the responder need not get the payoff closest to his reservation utility in ultimatum games. But, the argument is totally different from that developed in Section 5.1. It is the analogy with the fact that an outrageous offer is rejected in bargaining games that deters the proposer in the ultimatum game from making such an offer. It should be noted that the responder does not behave in the same way in the bargaining and ultimatum games. Thus, players do not confound the Rubinstein and ultimatum games. That is, using the automaton theory jargon players use different states of their machine/strategy to play the Rubinstein and ultimatum games. This should be contrasted with Samuelson [31] who develops an approach (in the automaton tradition) to explain

[30] The perturbed behavior that sustains such a belief is that the offer $(3/4, 1/4)$ (resp. $(1/4, 3/4)$) is proposed with the same probability in both R and U_1 (resp. U_2).

that rarely played games (presumably ultimatum games) may be lumped together with more frequently played games (presumably bargaining game) into a single state in order to save on the complexity costs that result from using strategies/machines that distinguish between the various games.

6. On endogenizing the analogy grouping

So far the analogy grouping used by the players was viewed as exogenous. Understanding how players categorize contingencies into analogy classes is a very challenging task left for future research. This short section is an attempt to suggest some tentative avenues for endogenizing the analogy grouping. More work both on the theory and empirical/experimental sides will be required to assess the relevance of these various approaches (and others).

6.1. Analogy-based expectation and similar play

One approach adopts the view that in order for player i to pool several nodes (j, h) into a single class of analogy, player i should himself consider playing in the same way in some pool of nodes. One difficulty is that in general player i need not move in the same nodes as player j, and therefore one should also worry about which nodes $h' \in H_i$ player i considers as being similar to nodes $h \in H_j$.

A class of situations in which this issue can be addressed simply is one in which whenever player i bundles two elements (j, h) and (j', h') into the same analogy class α_i, player i also has to move in h and h'. And the property is that player i behaves in the same way in nodes h and h'. An application of this idea is now being considered to illustrate how equilibria other than the subgame perfect Nash equilibrium may emerge.

Consider the following two-stage two-player game. Player 1 moves first and chooses between the normal form game G or G'. In both G and G', players 1 and 2 move simultaneously, and in both G and G', player 1 chooses in $A_1 = \{U, D\}$, player 2 chooses in $A_2 = \{L, R\}$. We assume that U is a dominant strategy in both G and G' for player 1. Player 2's best-response to U is R in game G, whereas it is L in game G'. Finally, we assume that player 1 derives a higher payoff when (U, R) is played in game G than when (U, L) is played in game G'. Finally, we assume that player 1 derives a higher payoff when (U, L) is played in game G' than when (U, L) is played in game G.

The unique subgame perfect Nash equilibrium is such that player 1 chooses game G and then (U, R) occurs.[31]

Suppose that player 1 puts in the same analogy class $(2, G)$ and $(2, G')$ in order to predict player 2's behavior. Note first that player 1 behaves in the same way in G and G' (he has the same dominant strategy in both games). Thus, the required property is satisfied. Second, it is readily verified that an equilibrium outcome in this analogy setting is that player 1 chooses G' (expecting player 2 to play L in both G and G'), since player 1 prefers (U, L) in game G' to (U, L) in game G.

[31] If he were to choose G' he would get the payoff attached to (U, L) in G', which is smaller.

6.2. The Nash equilibrium approach

Another approach is to view the analogy grouping as resulting from a learning/evolutionary process in which players would eventually learn to group contingencies in an optimal manner. There are several ways to implement this idea. Suppose, for example, that after choosing their analogy grouping, the same set of players keep interacting with each other until the system stabilizes to an analogy-based expectation equilibrium (for the chosen profile of analogy grouping), and suppose players have an opportunity to adjust their choice of analogy grouping in an optimal manner. From a game theoretic viewpoint, this amounts to considering a two stage-game: in the first stage, players choose their analogy partition, and in the second stage they play the resulting analogy-based expectation equilibrium. We assume that a player assesses his choice of analogy grouping according to the payoff that results from it in the second stage.[32]

It is beyond the scope of this paper to analyze the implications of this approach, but note that even if there are no costs to having extra analogy classes, the approach need not lead the players to select the finest analogy grouping.

As an illustration, it can be checked that in the example considered in Section 2, player 1 choosing the coarse partition and player 2 choosing the fine partition is part of an equilibrium.[33]

As a further illustration, consider the centipede game studied in Section 4, and assume that condition (3) holds. Focussing on the highest Pass rate equilibrium identified in Proposition 4, we observe that player 2 (resp. player 1) can guarantee a payoff of b_1 (resp. a_2) by choosing the coarsest partition (whatever the partition chosen by the other player). Thus, in the cooperative version of the centipede game where, say $a_2 > a_5$ and $b_1 > b_4$, (and for K not too small) the resulting equilibrium outcome must be that players get a payoff corresponding to a very high Pass rate. By contrast, in the dollar game or decreasing loser's payoff specifications however, both players choosing the finest partition constitutes an equilibrium.[34,35]

[32] There are alternative criteria. One alternative criterion is that players have aspiration levels and feel happy with their analogy partition, as long as the resulting payoff is not below their aspiration level. Another view is that players assess their analogy partitions according to the payoffs they perceive they will get in the ensuing equilibrium where the perception is assumed to be based on the ensuing analogy-based expectations (rather than on the effective payoff).

[33] (1) If both players use the fine partition, (U, R) is played in G and (D, R') in G' resulting in a payoff profile of $(2, 4)$.

(2) If player 1 uses the coarse partition and player 2 uses the fine partition, we have seen that (D, L) is played in G and (U, M) is played in G' resulting in an expected payoff profile of $(4, 2.5)$.

(3) If both players use the coarse partition, it can be checked that (D, L) in G and (U, R') in G' constitute an equilibrium resulting in an expected payoff profile of $(2.5, 2)$.

[34] Under those payoff specifications, it is quite essential not to be the loser (the player who does not Take). This in turn implies that it is a weakly dominant strategy choice to use the finest analogy grouping.

[35] The contrast between the implications of the various payoff specifications would also arise in the complexity approach developed by Rubinstein [30, Chapter 9].

7. Related literature

The literature on learning in games has emphasized that a Nash equilibrium need not emerge as a result of learning if, at the learning stage, players imperfectly observe the strategy of their opponents, and it has proposed the alternative concept of self-confirming equilibrium or conjectural equilibrium (see, in particular, Fudenberg and Kreps [8], Battigalli et al. [2] and Fudenberg and Levine [9, Chapter 6]). Most of this literature assumes that the learning agents observe the equilibrium path after each round of the learning process. A notable exception is Battigalli et al. [2] who consider more general signal structures. When general signal structures are considered, the main insights are about finding conditions under which conjectural equilibria are equivalent to Nash equilibria.[36] The analogy-based expectation equilibrium can be interpreted[37] as an equilibrium in this tradition in which each player i would not observe the equilibrium path, but only the average behavior of his opponents in every analogy class α_i at every step of the learning process.[38] But, by contrast with the insights developed in the conjectural equilibrium literature with general signal structures, the extra structure imposed by the analogy grouping idea has led us to propose a natural (or focal) notion of equilibrium, which happens to radically differ from existing equilibrium concepts (in particular, the Nash equilibrium).

Another relevant strand of literature is, of course, the literature on bounded rationality in games. Other approaches to bounded rationality (following the lead of Simon [34]) include the ε-equilibrium [24], the quantal response equilibrium [18], limited foresight models [12,13,15], games with procedurally rational players [22], and more recently the (partially) cursed equilibrium [7].[39]

The ε-equilibrium and the quantal response equilibrium approaches do not challenge the cognitive rationality of the players, since players are assumed (in equilibrium) to know perfectly the reaction function of their opponents. Like this paper, models of limited foresight do challenge the cognitive rationality of the players, but the implication of limited foresight is very different from that of analogy grouping.[40] In Osborne and Rubinstein [22] too, players do not rightly perceive the behavior of their opponent (players use a heuristic based on the idea that they play against Nature). But, unlike this paper, the interpretation is that players have an erroneous perception of the game being played (see Camerer [3] for an experimental account of misperceptions of games). Finally, Eyster–Rabin's [7] (partially) cursed

[36] The reason is that signals in such a general approach have no structures, and the (only) focal conjectural equilibria are the Nash equilibria.

[37] It may be more natural though to say that player i observes more, but that he only keeps track of such statistics because he thinks—possibly wrongly—that these are enough.

[38] Strictly speaking, one needs to have populations of players playing the game and the signals must bear on the average behavior throughout the population. Moreover, the notion of rationality is sequential in the analogy-based expectation equilibrium approach while it is not in the conjectural equilibrium approach.

[39] This review does not include approaches with fairness considerations or other approaches consisting in modifying the underlying preferences of the players.

[40] For example, limited foresight cannot explain cooperation in the finitely repeated prisoner's dilemma or the centipede game.

equilibrium is meant to capture common value situations in which players are imperfectly aware of the common value element of the game. Despite some common motivations, the approaches to partial sophistication (the main theme of this and their paper) are completely different in Eyster–Rabin and this paper.[41] Besides, Eyster–Rabin' setup is a static one of incomplete information whereas this paper considers multi-stage games with complete information.

There are many facets to analogy thinking. Other approaches in economics include the axiomatic approaches of Rubinstein [28] and Gilboa–Schmeidler [11] about similarity and case-based decision theory, respectively (which derive representation theorems for some axiomatic). These also include the automata theory developed for game theory by Rubinstein [27], and Abreu–Rubinstein [1] (see also Samuelson [31]).[42] It should be noted that none of these other approaches considers the treatment of expectations (as opposed to behaviors) by analogy. In a recent paper, Eliaz [6] proposes a solution concept in which players care about having simple representations of their opponents' strategies. In equilibrium, he imposes that the ''simplified'' representation held by every player i about player j's strategy coincides with player j's effective strategy (see also Spiegler [35] for a related solution concept with a similar feature). This feature is markedly different from the spirit of the analogy-based expectation equilibrium in which the representation of others' strategies need not coincide with their effective strategies.

Acknowledgments

Part of this research was done while I was visiting the Institute for Advanced Study the hospitality of which is gratefully acknowledged. I would like to thank P. Battigalli, O. Compte, E. Dekel, D. Ettinger, I. Gilboa, F. Koessler, B. MacLeod, E. Maskin, G. Nöldeke, R. Radner, A. Rubinstein, L. Samuelson, R. Spiegler, J. Tirole as well as the associate editor and referee for many helpful comments made at various stages of this research. I have also benefited from the comments made in seminars at Princeton University, Bonn University, Paris (ENS), NYU, Penn, the Institute for Advanced Study (Princeton), Yale, the ATT Labs, Champaign, Wisconsin, Rochester, Northwestern, UCL, University of Alicante, Ecole Polytechnique and Pompeu Fabra.

References

[1] D. Abreu, A. Rubinstein, The structure of Nash equilibrium in repeated games with finite automata, Econometrica 56 (1988) 1259–1282.

[2] P. Battigalli, M. Gilli, M.C. Molinari, Learning and convergence to equilibrium in repeated strategic interactions: an introductory survey, Ricerche Econ. 46 (1992) 335–378.

[41] The partially cursed equilibrium of Eyster–Rabin is also hard to interpret from a learning perspective.

[42] In the automaton setup, two different histories h and h' may induce the same state of player i's machine, and thus the same action of player i; Player i then acts in an analogous way in h and h'.

[3] C. Camerer, Mental representations of games, Mimeo, 1998.
[4] J. Dow, Search decisions with limited memory, Rev. Econ. Stud. 58 (1991) 1–14.
[5] U. Dulleck, J. Oechssler, The absent-minded centipede, Econ. Letters 55 (1997) 309–315.
[6] K. Eliaz, Nash equilibrium when players account for the complexity of their forecasts, Games Econ. Behav. 44 (2003) 286–310.
[7] E. Eyster, M. Rabin, Cursed Equilibrium, WP UC, Berkeley, 2002.
[8] D. Fudenberg, D. Kreps, Learning, experimentation and equilibrium in games, Mimeo Stanford University, 1988.
[9] D. Fudenberg, D. Levine, The Theory of Learning in Games, MIT Press, Cambridge, MA, 1998.
[10] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.
[11] I. Gilboa, D. Schmeidler, Case-based decision theory, Quart. J. Econ. 110 (1995) 605–639.
[12] P. Jehiel, Limited horizon forecast in repeated alternate games, J. Econ. Theory 67 (1995) 497–519.
[13] P. Jehiel, Learning to play limited forecast equilibria, Games Econ. Behav. 22 (1998) 274–298.
[14] P. Jehiel, Analogy-Based Expectation Equilibrium, WP Institute for Advanced Study, Princeton, 2001.
[15] P. Jehiel, Limited foresight may force cooperation, Rev. Econ. Stud. 68 (2001) 369–391.
[16] D. Kreps, P. Milgrom, J. Roberts, R. Wilson, Rational cooperation in the finitely repeated prisoners dilemma, J. Econ. Theory 27 (1982) 245–252.
[17] D. Kreps, R. Wilson, Sequential equilibria, Econometrica 50 (1982) 863–894.
[18] R. Mc Kelvey, T. Palfrey, An experimental study of the centipede game, Econometrica 60 (1992) 803–836.
[19] R. Mc Kelvey, T. Palfrey, Quantal response equilibrium for normal form games, Games Econ. Behav. 10 (1995) 6–38.
[20] R. Nagel, F. Tang, Experimental results on the centipede game in normal form: an investigation on learning, J. Math. Psychology 42 (1998) 356–384.
[21] A. Neyman, Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma, Econ. Letters 19 (1986) 227–229.
[22] M. Osborne, A. Rubinstein, Games with procedurally rational players, Amer. Econ. Rev. 88 (1998) 834–847.
[23] M. Piccione, A. Rubinstein, On the interpretation of decision problems with imperfect recall, Games Econ. Behav. 20 (1997) 3–24.
[24] R. Radner, Can bounded rationality resolve the prisoner's dilemma?, in: A. Mas-Colell, W. Hildenbrand (Eds.), Essays in Honor of Gerard Debreu, Amsterdam, North Holland, 1986, pp. 387–399.
[25] P. Reny, Common belief and the theory of games with perfect information, J. Econ. Theory 59 (1993) 257–274.
[26] R. Rosenthal, Games of perfect information, predatory pricing and the chain-store paradox, J. Econ. Theory 25 (1982) 83–96.
[27] A. Rubinstein, Finite automata play a repeated prisoner's dilemma, J. Econ. Theory 39 (1986) 145–153.
[28] A. Rubinstein, Similarity and decision-making under risk, J. Econ. Theory 46 (1988) 145–153.
[29] A. Rubinstein, On price recognition and computational complexity in a monopolistic model, J. Polit. Economy 101 (1993) 473–484.
[30] A. Rubinstein, Modeling Bounded Rationality, MIT Press, Cambridge, MA, 1998.
[31] L. Samuelson, Analogies, adaptation, and anomalies, J. Econ. Theory 97 (2001) 320–366.
[32] R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, Int. J. Game Theory 4 (1975) 22–55.
[33] R. Selten, R. Stoecker, End behavior in sequences of finite prisoner's dilemma supergames, J. Econ. Behav. Organ. 7 (1986) 47–70.
[34] H. Simon, A behavioral model of rational choice, Quart. J. Econ. 69 (1955) 99–118.
[35] R. Spiegler, Simplicity of beliefs and delay tactics in a concession games, Games Econ. Behav. 47 (2004) 200–220.