Analogy-Based Expectation Equilibrium*

Philippe Jehiel†

March 2001

Abstract

It is assumed that players bundle nodes in which other players must move into analogy classes, and players only have expectations about the average behavior in every class. A solution concept is proposed for multi-stage games with perfect information: at every node players choose best-responses to their analogy-based expectations, and expectations are correct on average over those various nodes pooled together into the same analogy classes. The approach is applied to a variety of games. It is shown that a player may benefit from having a coarse analogy partitioning. And for simple analogy partitioning, (1) initial cooperation followed by an end opportunistic behavior may emerge in the finitely repeated prisoner's dilemma (or in the centipede game), (2) an agreement need not be reached immediately in bargaining games with complete information.

Key words: Game theory, bounded rationality, reasoning by analogy.

JEL numbers: C72, D81.

*I would like to thank O. Compte, D. Ettinger, I. Gilboa, B. MacLeod, E. Maskin, G. Nöldeke, R. Radner, A. Rubinstein, L. Samuelson, R. Spiegler, and E. Dekel and three anonymous referees for many helpful comments made at various stages of this research. I have also benefitted from the comments made in seminars at Princeton University, Bonn University, Paris (ENS), NYU, Penn, the Institute for Advanced Study (Princeton), Yale, and the ATT Labs.

† CERAS, Paris and UCL, London. mailing address: C.E.R.A.S.-E.N.P.C., C.N.R.S. (URA 2036), 28 rue des Saints-Pères, 75007 Paris, France; e-mail: jehiel@enpc.fr.

1 Introduction

Received 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 in which the standard approach is inappropriate. In chess, it is clearly impossible to know (learn) what the opponent might do for every board position.

In this paper, we investigate situations in which players form their expectations about others' behavior by *analogy* between *several* contingencies as opposed to for every *single* contingency in which each of these other players must move.[1] More precisely, each player i *bundles* nodes at which players other than i must move - a bundle is called an *analogy class*. And player i only forms expectations about the *average* behavior in each analogy class that he considers.

In other words, player i is viewed here as simplifying what he wants to know (learn) about others' behavior:[2] Player i categorizes nodes in which other players must move into analogy classes. And only the average behavior in each analogy class is being considered by player i.

We use the word 'analogy' because in two nodes belonging to a same class, the expectation formed by the player is the same. Besides, the equilibrium expectation in an analogy class will be assumed to coincide with the effective average behavior in the class. Accordingly, nodes which are visited more often will contribute more to the expectation, and the behaviors in those nodes will contaminate the expectation used in all nodes 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 - we believe - a key feature of the analogy idea.[3]

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. 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.

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.[4]

[1] This approach seems particularly appropriate in situations with many contingencies (like chess) so that learning behavior for every possible contingency seems too hard (impossible).

[2] This makes learning easier and successful learning more plausible.

[3] It should be noted that what is considered here is the idea of forming expectations by analogy as opposed to acting by analogy (see discussion section).

[4] Extensions to incomplete information setups raise no conceptual difficulties, but make the exposition

The partitioning into analogy classes used by the players is given exogenously.[5] It is viewed as part of the description of the strategic environment. 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. We require that 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 identically labelled.[6]

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.[7] Second, for each player i and analogy class α_i, player i's expectation $\beta_i(\alpha_i)$ is *consistent* with the average behavior in class α_i as induced by the strategy profile σ (where the behavior of player j in 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).[8]

Clearly, if 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. However, when at least one player does not use the finest partitioning, the play of an analogy-based expectation equilibrium will in general differ from that of a Subgame Perfect Nash equilibrium (or even from that of a Nash equilibrium). We also note that in finite environments an analogy-based expectation equilibrium always exists.

In the second part of the paper, we investigate a few properties of analogy-based expectation equilibria in a variety of games. We first observe that sometimes a player may *benefit* from having a coarse analogy partitioning as compared with the finest partitioning.

---

notationally heavy.

[5] One might think of the partitioning as resulting from the past experiences of the players and also from the way the strategic interaction is framed to the players thus triggering some connections with past experiences (the so called framing effect, see Tversky-Khaneman 1981).

[6] Strictly speaking, it is enough to require that there is a bijection between the two action spaces. Note also that our formalism allows for analogies accross different players.

[7] More precisely, player i's strategy σ_i is a best-response (after 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.

[8] We think of the consistency requirement 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). And 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.)

Clearly, this is not so if this player plays against Nature or if other players have a dominant strategy. Then a coarse partitioning has the sole effect of making this player 's choice of strategy possibly suboptimal without affecting the behaviors of others. But, otherwise, a coarse partitioning of, say, player i may well induce (in equilibrium) a change of strategies of players other than i (as a response to a change of strategy of player i). When such a change of strategies is good for player i, player i may in equilibrium end up with a strictly higher payoff.

We next apply the analogy-based expectation approach to the so called finite horizon paradoxes. For simple analogy partitioning, we show both in the centipede game and in the finitely repeated prisoner's dilemma that there may be equilibria in which there is a fair amount of cooperation throughout the game except possibly toward reaching the end of the game at which time some opportunistic behavior may occur.

To illustrate the claim, consider a variant of the finitely repeated prisoner's dilemma in which there are many periods, there is no discounting and the exact values of the stage game prisoner's dilemma payoffs are independently drawn from period to period according to some pre-specified distribution (with finite support). And assume that both players categorize histories into two analogy classes according to whether or not some opportunistic behavior was previously observed (within the game).

Playing cooperatively most of the time except if some opportunistic behavior previously occurred or toward the end of the game (if the immediate gain from switching to an opportunistic behavior is sufficiently high) is part of an analogy-based expectation equilibrium.

To see this, consider the expectations induced by the behaviors just described. Each player should *expect* the other player (1) to play opportunistically whenever some opportunistic behavior previously occurred and (2) to play cooperatively (on average) with a large probability otherwise (if the number of repetitions is large). Given such expectations, playing opportunistically is optimal whenever some opportunistic behavior previously occurred. And, when no opportunistic behavior previously occurred, playing cooperatively in all but a few periods toward the end is also optimal because players perceive that by playing opportunistically they will trigger a non-cooperative phase whereas by playing cooperatively they expect the other player to continue playing cooperatively with a large probability.

The key reason why the logic of backward induction fails here is that players do not perceive exactly *when* the other player will start having an opportunistic behavior. As a result of this fuzzy perception (which is due to their analogy partitioning), players play cooperatively most of the time because on average it is true that by playing cooperatively the other player keeps playing cooperatively with a large probability.

It should be noted that players do consider playing opportunistically toward reaching the end of the game, even if no opportunistic behavior previously occurred. This is so whenever the immediate gain from playing opportunistically offsets the cost of triggering a non-cooperative phase till the end of the game (as opposed to maintaining the cooperative phase). In this sense, players do perceive the time structure of the interaction even though they do not perceive the exact time structure of the strategy employed by their opponent.[9]

We also briefly consider the infinitely repeated prisoner's dilemma. We observe that strategy profiles in which some deviations are not punished can be sustained with the analogy-based expectation approach. The point is that, while such a deviation would be profitable, it need not perceived as such if the corresponding node is bundled with nodes in which there would be an effective punishment. As a result (of such an analogy partitioning), the involved player perceives an average punishment, which deters him from deviating. Thus, in repeated games, the analogy-based approach permits less systematic punishments than the standard approach does.

Our next application deals with ultimatum and bargaining games. Suppose that players can make any possible offer, but that they have expectations about the acceptance probability only according to whether the offer is above or below a threshold (i.e., whether or not their offer is generous). We show that (1) the responder in a take-it-or-leave-it offer game may get a payoff that lies strictly above his reservation utility (i.e. his payoff from refusing any agreement), (2) there may be no immediate agreement in a (complete information) bargaining game in which players alternate making offers.

The effect of analogy reasoning here is to reduce the set of offers that players consider in equilibrium. If a player makes a generous offer, he will always consider the least generous offers among these. This is because (due to his analogy partitioning) he has the same (acceptance) expectation for all such offers, and the least generous offer among these is clearly the one he likes best given such an expectation. The analysis of ultimatum and bargaining games follows.

In the last part of the paper, we provide some general discussion. We first differentiate the analogy-based expectation equilibrium from other solution concepts, in particular related to the idea of imperfect recall (and of imperfect information).[10] Second, we suggest two principles that may help structure analogy partitioning. The first principle applies to those games in which all players must move in the same nodes, and we consider the extra requirement that a player should himself behave in the same way in all nodes associated with the same

[9] The analogy approach thus permits an endogenous treatment of the end effect identified in experiments (see Selten-Stocker 1986).

[10] An alternative interpretation for the concept is also proposed and discussed.

analogy class. The second principle is that all analogy classes considered by every player should be reached with positive probability along the played path. For both principles, we provide examples in which the prediction of the analogy-based approach is in sharp contrast with the conventional approach. We also analyze the issue of multiplicity of equilibria, and we discuss some of the related literature on bounded rationality.

2 A general framework

2.1 The class of games

We consider multi-stage games with almost perfect information and perfect recall. That is, in each stage every player knows all the actions that were taken at any previous stage (including those exogenous events determined by Nature at any previous stage), and no information set contained in the current stage provides any knowledge of play in that stage.[11]

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.[12]

The standard representation of an extensive form game in class Γ includes the set of players $i = 1, ...n$ denoted by N, the game tree Υ (specifying who moves when and over which space, including the exogenous events chosen by Nature), and the preferences $\succsim_i$ of every player i over outcomes in the game.

A node in the game tree Υ will be denoted by h; it contains information about all the actions, including those by Nature, that were taken at any stage prior to node h. The set of nodes h will be denoted by H. The set of nodes at which player i must move will be denoted by H_i. For every node $h \in H_i$, we let $A_i(h)$ denote player i's action space at node h.

Remark: When interpreting experiments, it may be meaningful to view the players as being engaged in a variety of games as opposed to only one game.[13] One can represent this as a metagame made of an extra move by Nature in stage 0 which would determine the effective game to be played (according to the frequency with which each (original) game was played).

[11] Also, simultaneous moves are allowed, but each player moves at most once within a given stage.

[12] In some applications, we will consider infinite action spaces and/or infinitely many stages. The solution concept will easily generalize to these applications.

[13] For example, bargaining and ultimatum games or centipede games of various lengths...

Classes of analogy:

Each player i forms an expectation about the behavior of other players j, $j \neq i$. However, player i does not form an expectation about every player j's behavior in every contingency $h \in H_j$ in which player j must move. He pools together several contingencies in which other players must move, and he forms an expectation about the *average* behavior in these pooled contingencies. Such a *pool* is referred to as a *class of analogy*.

Formally, each player i partitions the set $\{(j,h) \in N \times H_j,\ j \neq i\}$ into subsets α_i referred to as analogy classes.[14] 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.[15] 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.

Remark: At first glance, there is some resemblance between an analogy class and an information set in an extensive form game with incomplete information. However, note that An_i refers to a partitioning of the nodes where players *other than* i must move (as opposed to a partitioning of the nodes where player i *himself* must move as in the notion of player i's information set).[16]

Strategic environment:

A strategic environment in our setup not only specifies the set of players N, the game tree Υ and players' preferences $\succsim_i$. It also specifies how the various players partition the set of nodes at which other players must move into classes of analogy, which is summarized in An. A strategic environment is thus formally given by $(N, \Upsilon, \succsim_i, An)$.

2.2 Concepts

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

[14] 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 $k \neq k'$.

[15] More generally, we could allow the players to relabel the original actions of the various players as they wish. From that prespective, $A_j(h)$ should only be required to be in bijection with $A_{j'}(h')$ (as opposed to being equal). Describing this and the subsequent notion of consistency would require heavy notations without adding anything to the concept. It is therefore ignored for expositional reasons.

[16] We will offer more discussion throughout the paper on the relationship between analogy reasoning and incomplete information (and imperfect recall) in extensive form games.

average behavior in class α_i.

Remark: Note again the different nature of $\beta_i(\cdot)$ 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.)

Strategy:

A behavior strategy for player i 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.[17]

Formally, a behavior strategy for player i is denoted by σ_i. 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 also let σ_{-i} denote the strategy profile of players other than i, and we let σ denote the strategy profile of all players.

Sequential rationality:

The criterion used by the players to choose their strategies given their analogy-based expectations is as follows. 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 most natural strategy profile compatible with player i's partial expectation β_i.[18]) And the criterion considered by player i is that of best-response against this induced strategy profile after every node where player i must move.

Formally, for every β_i and $j \neq i$, we define the β_i-perceived strategy of player j, $\sigma_j^{\beta_i}$, as[19]

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

Given player i's strategy σ_i and given node h, we let $\sigma_i \mid_h$ denote the continuation strategy of player i induced by σ_i from node h onwards. Similarly, we let $\sigma_{-i} \mid_h$ and $\sigma \mid_h$ be the strategy profiles induced by σ_{-i} and σ, respectively, from node h onwards. We also let $u_i^h(\sigma_i \mid_h, \sigma_{-i} \mid_h)$ denote the expected payoff obtained by player i when the play has reached node h, and players behave according to the strategy profile σ.[20]

[17] Mixed strategies and behavior strategies are equivalent since we consider games of perfect recall. The behavior strategy formulation is better suited to define the consistency condition (see below).

[18] If for every h such that $(j, h) \in \alpha_i$, the behavior strategy of player j at node h is given by $\beta_i(\alpha_i)$, then the average of these (whatever the weighting of the various elements of α_i) must be $\beta_i(\alpha_i)$. A richer setup would allow player i to consider any strategy profile for players other than i that is compatible with his partial knowledge β_i (see Remark 2 after the definition of analogy-based expectation equilibria).

[19] This defines a strategy for player j because all (j, h) where $h \in H_j$ belong to one and only one α_i since the set of α_i is a partition of $\{(j, h) \in N \times H_j,\ j \neq i\}$.

[20] These functions can formally be derived from $\succsim_i$ and the distributions over outcomes induced by $\sigma \mid_h$.

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 $h \in H_i$,*

$$u_i^h(\sigma_i \mid_h, \sigma_{-i}^{\beta_i} \mid_h) \geq u_i^h(\sigma_i' \mid_h, \sigma_{-i}^{\beta_i} \mid_h).$$

Consistency:

In equilibrium, we 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.

We 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-Levine 1998), we distinguish according to whether or not consistency is also required for analogy classes that are reached with probability 0 in equilibrium.[21]

To present formally the consistency idea, we denote by $P^\sigma(h)$ 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 $\alpha_i \in An_i$:*

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

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

This definition deserves a few comments. The view is that each player i happens to make *consistent* (or correct) analogy-based expectations as a result of *learning*. Suppose 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 node (j,h) is visited (according to σ) relative to

[21] 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.

other elements in α_i. The correct weighting of $\sigma_j(h)$ should thus be $P^\sigma(h)/\left(\sum_{(j,h)\in\alpha_i} P^\sigma(h)\right)$, 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) that places restrictions also on those expectations.

Formally, we define Σ^0 to 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 (thus implying that $\sigma_j(h)$ has full support on $A_j(h)$ for all j, $h \in H_j$).

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 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.*

Solution concepts:

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

Definition 4 *An assessment (σ, β) is 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* consistent *with σ.*

Definition 5 *An assessment (σ, β) is an 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 *with σ.*

Remark 1*:* To the extent that the number of analogy classes α_i considered by player i is small, player i has few features of the other players' behavior to learn, which makes the consistency requirement more plausible from a learning perspective than in the perfect rationality paradigm.

Remark 2*:* A priori there are strategies other than $\sigma_{-i}^{\beta_i}$ that could generate the analogy-based expectation β_i.[22] A more elaborate criterion than the one considered in Definition 1 would view player i as playing a best-response against some strategy profile σ'_{-i} compatible with β_i but not necessarily $\sigma_{-i}^{\beta_i}$. The corresponding solution concepts would be somewhat more complicated to present (but most of the insights developed below would continue to hold for such alternative specifications).

Remark 3*:* We have assumed that player i's analogy classes are partitions of the nodes where players *other than* i must move. In some cases, it may be meaningful to allow players to predict the behavior of other players also based on their *own* behavior. There is no difficulty with allowing the analogy classes α_i to also include nodes (i, h) such that at node h player i must choose an action in $A(\alpha_i)$ (the same action space as the one faced by the other players involved in α_i). However, it should be understood that the corresponding analogy-based expectation $\beta_i(\alpha_i)$ is used by player i only to construct a strategy profile for players other than i (see Definition 1).[23]

Remark 4*:* The setup could easily be extended to cover the case where players have private information. However, this would significantly complicate the description of the setup. For expositional (rather than conceptual) reasons, we have chosen to focus on games with almost perfect information.

2.3 Preliminary results

We conclude this general presentation by making two simple observations. The first one shows the relation to subgame perfection 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.

Formally, 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* $(j', h') \in \alpha_i$. We have:

Proposition 1 *Consider an environment* $(N, \Upsilon, \%_i, An)$ *in which all players use the finest*

[22] In general (except for $\sigma_{-i}^{\beta_i}$), to check that σ'_{-i} generates β_i it is indispensable to know the frequency of visits of every node $h \in \alpha_i$ (as given by the candidate strategy profile σ).

[23] We have chosen not to present the concept with that extension because it could create an extra source of confusion (with the notion of information set).

analogy partitioning. Then if (σ, β) *is an analogy-based expectation equilibrium of* $(N, \Upsilon, \succsim_i, An)$*,* σ *is a Subgame Perfect Nash Equilibrium of* $(N, \Upsilon, \succsim_i)$*.*

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

Remark: 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 $\sigma_{-i}^{\beta_i}$ (after every node h). But, $\sigma_{-i}^{\beta_i}$ need not (in general) coincide with σ_{-i} as in a Subgame Perfect Nash equilibrium. This will be further illustrated throughout the paper.

Proposition 2 *(Existence) Every finite environment* $(N, \Upsilon, \succsim_i, An)$ *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 (Kreps and Wilson 1982). We mention the argument, but for space reasons we 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 1975).[24] 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. ■

3 Various effects of analogy reasoning

3.1 Analogy reasoning can be good or bad

We wish to illustrate that bundling contingencies by analogy can either benefit or hurt a player. To this end, we consider the following environment. Two normal form games G and G' are being played in parallel. Game G is played with probability ν and game G' is played with probability $1-\nu$. (In the formulation of Section 2, the game tree Υ consists of a first move by Nature about the selection of the game - according to the probabilities ν and $1-\nu$ - then followed by the normal form game G or G' accordingly.) There are two players $i = 1, 2$ in G and G'. In both G and G', player i must choose an action a_i in the *same* finite action space A_i.

[24]This requires amending Definition 1 to incorportate such constraints in the maximization programmes.

In the game tree Υ, a node can be identified with a normal form game G or G'. We assume that player 2 uses the finest partitioning (i.e., player 2 uses two analogy classes $\{(1, G)\}$ and $\{(1, G')\}$).

We wish to compare the equilibrium payoff obtained by player 1 in each of the subgames G, G' according to whether player 1 uses the *finest* partitioning or the *coarsest* partitioning (in the latter case player 1 pools together the two subgames G and G' into a single class of analogy $\{(2, G), (2, G')\}$).

Claim 1: Suppose player 2 has a dominant strategy[25] in both games G and G'. Player 1's equilibrium payoff - in both G and G' - is no smaller when player 1 uses the finest partitioning as opposed to the coarsest partitioning.

Proof. Whatever the partitioning of player 1, player 2 will in equilibrium select his dominant strategy in both G and G'. The finest partitioning of player 1 allows player 1 to pick a best-response to player 2's dominant strategy in both G and G', which is the highest payoff player 1 can hope to get (in both G and G') given player 2's behavior. ■

Within the context of Claim 1, it is immediate to construct an example in which player 1 's equilibrium payoff is strictly lower when he uses the coarse partitioning as opposed to the finest partitioning. (Such an example must be such that player 2 's dominant strategy is not the same in games G and G' , and thus player 1's analogy-based expectation is not accurate for games G and G' in isolation.)

When player 2 has no dominant strategy, however, analogy reasoning may benefit player 1, as the following example shows.

Example 1: Consider the following situation

	L	M	R
U	$5,1$	$0,1$	$2,2$
D	$3,1$	$3,0$	$1,0$

Game G

	L	M	R
U	$3,0$	$3,1$	$1,0$
D	$0,1$	$5,1$	$2,2$

Game G'

where in each cell the left and right numbers indicate players 1 and 2's payoffs, respectively. Both games are assumed to be played with equal probability, i.e. $\nu = \frac{1}{2}$. In both G and G', the action space of players 1 and 2 are $A_1 = \{U, D\}$ and $A_2 = \{L, M, R\}$, respectively.

The example is such that both G and G' have a unique Nash equilibrium, which is UR in game G and DR in game G'. Thus, when both players use the finest partitioning, player 1 gets a payoff of 2 in both subgames.

Suppose now that player 1 uses the coarsest partitioning (while player 2 uses the finest). The following assessment is an analogy-based expectation equilibrium.

[25] This dominant strategy need not be the same in both games G and G'.

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 $\{(2,G),(2,G')\}$). Player 2 expects player 1 to play D in game G and U in game G'.

To check that the above assessment is an equilibrium, note that given the strategy profile, players' analogy-based expectations are consistent. Then 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'.

Finally, note that according to the above strategy profile player 1 gets a payoff of 3 in both G and G', which is strictly larger than 2 - the equilibrium payoff obtained by player 1 when he uses the finest partitioning. ■

The key feature of Example 1 is that player 2 does not play in the same way when player 1 uses the finest partitioning and when he uses the coarsest partitioning. It is still the case that the coarseness of player 1's partitioning induces player 1 not to optimize against player 2's behavior in G and G' (because the best-response would be U (and not D) in game G and D (and not U) in game G'). However, it allows player 1 to find it optimal to play D (resp. U) in game G (resp. G'), which in turn induces player 2 to play an action that is more favorable to player 1.

Remark 1*:* In the analogy-based expectation equilibrium shown in Example 1, both players 1 and 2 behave differently in games G and G'. Thus, even by varying the payoff matrix of games G and G', it is not possible to interpret the equilibrium outcome as emerging from an imperfect information (of either player) as to which game (G or G') is being played.

Remark 2*:* In Example 1, when player 1 uses the coarsest partitioning, there is also an equilibrium in which UR is played in game G and DR is played in game G' as in the finest partitioning case. Modify the specification of game G so that player 2 has a dominant strategy which is to play L. It can be checked that when player 1 uses the coarsest partitioning the assessment shown in Example 1 is the only analogy-based expectation equilibrium in this modified example. Thus, in this modified setup, player 1 benefits from the coarse partitioning in subgame G' whatever the equilibrium under consideration.

Remark 3*:* If one insists on having equilibria that employ pure strategies, player 2 should have at least three actions for an example of the sort displayed in Example 1 to work. Otherwise, similar conclusions can be derived with 2x2 games and mixed strategy equilibria.

Comment: In the above situation we have assumed that the same player 2 were to play in both games G and G'. Of course, an alternative interpretation is that the player other than 1 involved in game G is not the same as the one involved in game G', say player 2 in game G and player $2'$ in game G'. The partitioning of player 1 considered above corresponds then to $\{(2, G), (2', G')\}$. For that interpretation, it is essential to allow player 1 to treat by analogy nodes in which several *different* players (here players 2 and $2'$) are involved.

3.2 Centipede game

Consider the centipede game CP_K (first introduced by Rosenthal 1982) and depicted in Figure 1.

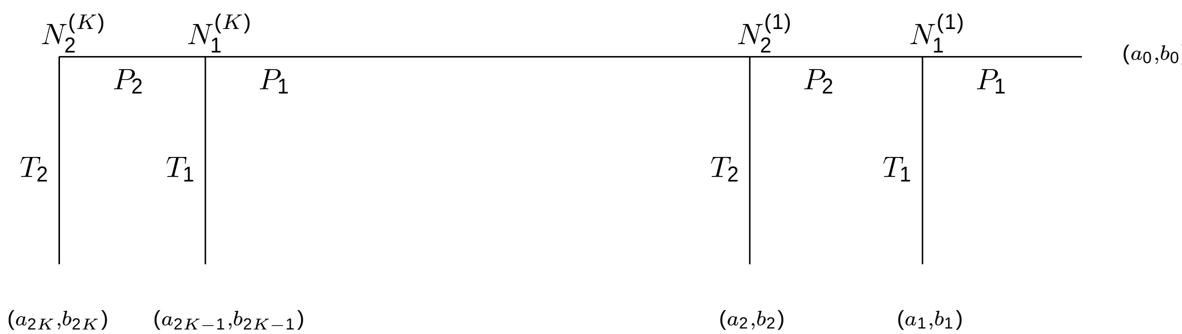

Figure 1: The centipede game

It is a $(2K)$-period extensive form game described as follows. There are two players $i = 1, 2$ who move in alternate order. In each period, the player whose turn it is to move, say player i, may either *Take* or *Pass*, i.e. $A_i = \{Pass, Take\}$.[26] If he Takes, this is the end of the game. If he Passes, the game proceeds to the next stage where it is the other player's turn to move unless the game has reached the last period $2K$ in which case this is the end of the game. Player 1 moves in the last period $2K$, player 2 moves in the last but one period and so on. Nodes at which player 1 must move are labelled $N_1^{(k)}$, $k = 1, ...K$ where $N_1^{(1)}$ designates the last node (i.e., period $2K$) at which player 1 must move, $N_1^{(2)}$ the last but one, and so on till $N_1^{(K)}$ the first node (i.e, period 2) at which player 1 must move. Similarly, nodes at which player 2 must move are labelled $N_2^{(k)}$, $k = 1, ...K$ where $N_2^{(1)}$ designates the last node (in period $2K - 1$) at which player 2 must move, and $N_2^{(K)}$ the first node (in period

[26] We implicitly assume in the following that the players label these actions the same way.

1) at which player 2 must move. If player 2 Takes at node $N_2^{(k)}$, the payoffs to players 1 and 2 are a_{2k} and b_{2k}, respectively. If player 1 Takes at node $N_1^{(k)}$, the payoffs to players 1 and 2 are a_{2k-1} and b_{2k-1}, respectively. If player 1 Passes at node $N_1^{(1)}$, the payoffs to players 1 and 2 are a_0 and b_0, respectively. All a_t and b_t, $t = 0, ...2K$ are assumed to be integers that satisfy for all $k \geq 1$:

$$\begin{aligned} a_{2k-1} &> a_{2k-2} > a_{2k+1} > a_{2k} \\ b_{2k-2} &> b_{2k-3} > b_{2k} > b_{2k-1} \end{aligned} \tag{2}$$

These conditions guarantee that (1) the unique Subgame Perfect Nash Equilibrium (SPNE) of CP_K is such that player 2 Takes in the first period (this follows from $a_{2k+1} > a_{2k}$ and $b_{2k} > b_{2k-1}$), and (2) in any period $t \leq 2K - 2$, both players are 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 $a_t > a_{t+2}$ and $b_t > b_{t+2}$ for all $t \leq 2K - 2$).

The prediction of the SPNE sounds relatively unintuitive, especially when the number of periods $2K$ is large (because then taking in the first node seems to induce very severe losses for both players). As we now illustrate, the analogy approach explains why players may Pass most of the time in the centipede game, at least for long enough versions of the game.

In order to deal with the effect of increasing the number of periods in CP_K, we will consider infinite sequences of integers $(a_k)_{k=0}^{\infty}$, $(b_k)_{k=0}^{\infty}$ satisfying (2). We will also assume that the difference between two consecutive payoffs is uniformly bounded from above. That is, there exists $\Delta > 0$ such that for all $t \geq 0$,

$$|a_t - a_{t+1}| < \Delta \text{ and } |b_t - b_{t+1}| < \Delta. \tag{3}$$

Regarding analogy partitioning, we will mostly consider the case in which both players use the coarsest partitions as their analogy device. That is, each player i is assumed to pool together *all* the nodes $N_j^{(k)}$ at which player j, $j \neq i$ must move into a single class of analogy α_i:

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

The strategic environment is thus described by the set of players $N = \{1, 2\}$, the game tree CP_K, players' preferences $\%_i$ as defined by a_t, b_t, and the analogy partitioning structure An as described by α_1 and α_2: $(N, CP_K, \%_i, An)$.

Player i's analogy-based expectation β_i reduces here to a single probability measure $\beta_i(\alpha_i) = \lambda^i \cdot Pass + (1 - \lambda^i) \cdot Take \in \Delta A_j$, which stands for player i's expectation about the average behavior of player j throughout the game.

We first consider assessments (σ, β) such that the strategy of every player is pure (i.e. for every i, $h \in H_i$, $\sigma_i(h) \in A_i$). And we consider the following condition:

$$\frac{K-1}{K}b_{2k} + \frac{1}{K}b_{2k+1} > b_{2k+2} \text{ for all } k \geq 0. \tag{4}$$

Proposition 3 *Suppose that condition (4) holds, and consider the environment* $(N, CP_K\ \%_i, An)$. *There are two possible equilibrium paths corresponding to self-confirming analogy-based expectation equilibria in pure strategies: Either player* 2 *Takes in the first period or the game reaches period* $2K$ *in which player* 1 *Takes.*

Proof. (a) We first prove that the two mentioned outcomes can be obtained as analogy-based expectation equilibrium outcomes.

(i) Observe first that the Subgame Perfect Nash Equilibrium outcome corresponds to the analogy-based expectation equilibrium (σ, β) in which for $i = 1, 2$, $\beta_i(\alpha_i) = Take$ and $\sigma_i(N_i^{(k)}) = Take$ for all $k = 1, ...K$.

(ii) Consider the strategy profile σ such that player 2 Passes always and player 1 Takes in the last period $2K$.

To be consistent with σ, the analogy-based expectation of player 1 must be that player 2 Passes with probability 1, i.e. $\beta_1(\alpha_1) = Pass$ (since player 2 Passes always when he has to move).

To be consistent with σ, the analogy-based expectation of player 2 must be that player 1 Passes with probability $\frac{K-1}{K}$, since (according to σ) each node $N_1^{(k)}$, $k = K, ...1$ is reached with probability 1, i.e. $P^\sigma(N_1^{(k)}) = 1$, (so that they have equal weighting), and player 1 Passes (with probability 1) in all nodes $N_1^{(k)}$, $k = K, ...2$ and Takes in node $N_1^{(1)}$. Thus, $\beta_2(\alpha_2) = \frac{K-1}{K}Pass + \frac{1}{K}Take$.

The (sequential) best-response of player 1 to the analogy-based expectation β_1 is to Take in the last node $N_1^{(1)}$. Thus, it is to play according to σ_1.

When condition (4) holds, the best-response of player 2 to the analogy-based expectation β_2 is to Pass always (since by Taking at $N_2^{(k+1)}$, player 2 would only get b_{2k+2}, which is less than the expected payoff he gets by Passing at $N_2^{(k+1)}$ and Taking at $N_2^{(k)}$, say, i.e. $\frac{K-1}{K}b_{2k} + \frac{1}{K}b_{2k+1} > b_{2k+2}$). Thus, it is to play according to σ_2.

Altogether the above considerations show that the assessment (σ, β) is an analogy-based expectation equilibrium.

(b) It remains to show that there are no other possible outcomes in any pure strategy self-confirming analogy-based expectation equilibrium. Observe first that if an outcome other than 'Player 2 Takes in the first period' emerges (as a self-confirming analogy-based expectation equilibrium outcome), it must correspond to an analogy-based expectation equilibrium outcome. (This is because the unique analogy class of every player is then reached with strictly positive probability.)

Consider the outcome in which player i Takes at node $N_i^{(k)}$, and $N_i^{(k)}$ differs from $N_2^{(K)}$. If a pure strategy analogy-based expectation equilibrium leads to that outcome, it must be (by consistency) that player i's analogy based expectation satisfies $\beta_i(\alpha_i) = Pass$, since on the equilibrium path, player j would always Pass. Player i's best response to such a β_i depends on whether $i = 1$ or 2. If $i = 1$, player 1's best response to $\beta_1(\alpha_1) = Pass$ is to Take at node $N_1^{(1)}$ (which corresponds to an outcome already identified as a possible analogy-based expectation equilibrium outcome). If $i = 2$, player 2's best response to $\beta_2(\alpha_2) = Pass$ is to Pass always, which is in contradiction with player 2 Taking at node $N_2^{(k)}$.

Finally, the outcome in which both players Pass in every period cannot be an analogy-based expectation equilibrium outcome because whatever player 1's expectation, player 1 strictly prefers Taking at node $N_1^{(1)}$ to Passing always. ■

Comment 1*:* The two outcomes mentioned in Proposition 3 remain equilibrium outcomes even if player 1 uses a partitioning other than the coarsest, as long as player 2 uses the coarsest partitioning.[27]

Comment 2*:* Consider the case where player 2 uses the coarsest partitioning and player 1 uses the finest partitioning (and condition (4) holds). As mentioned in Comment 1, Take by player 1 in the last node can be sustained as the equilibrium outcome of an analogy-based expectation equilibrium. Note that in this equilibrium, player 2 behaves in the same way in every node where he must move, which is to be related to his bundling of *all* nodes in which player 1 must move into a single class of analogy. We will suggest such a *principle* for refining analogy-based expectation equilibria in Section 4.

Proposition (3) leaves open what happens when condition (4) does not hold.[28] And it does not deal with equilibria in mixed strategies. The next Proposition provides the main missing information (still assuming that conditions (2) and (3) hold):

Proposition 4 *There exists an integer $\overline{m}$ such that for all $K > \overline{m}$: (1) $(N, CP_K, \%_i, An)$ has an analogy-based expectation equilibrium (σ, β) in which each player i Passes with probability*

[27] If one additionally requires that for all k, $b_{2k} < \frac{b_{2k-1}+b_{2k-2}}{2}$, then these are the only possible outcomes of self-confirming analogy-based expectation equilibria in pure strategies. (The point is that for a pure outcome other than that of the SPNE to emerge as a self-confirming analogy-based expectation equilibrium, it should be that $\lambda^2 \geq \frac{1}{2}$. But, then the best-response of player 2 to $\beta_2(\alpha_2) = \lambda^2 Pass + (1-\lambda^2) Take$ is to Pass always, thus leading to the wished conclusion.)

[28] Take at the last node may then fail to be the outcome of an analogy-based expectation equilibrium in pure strategies. This is, for example, the case when $\frac{K-1}{K} b_0 + \frac{1}{K} b_1 > b_2$ (because then player 2 would strictly prefer Taking in the last but one node rather than Passing always).

1 *in the first* $K-\overline{m}$ *nodes, i.e. in every* $N_i^{(k)}$, $k=K,...K-\overline{m}$. *(2) Any self-confirming analogy-based expectation equilibrium of* $(N, CP_K, \%_i, An)$ *in which each player i Passes with probability* 1 *in the first node* $N_i^{(K)}$ *is such that each player i Passes with probability* 1 *in the first* $K-\overline{m}$ *nodes, i.e. in every* $N_i^{(k)}$, $k=K,...K-\overline{m}$.

Proof. (1) Suppose $\beta_i(\alpha_i) = \lambda^i.Pass + (1-\lambda^i).Take$ with $\lambda^i \geq \frac{1}{2}$ for $i=1,2$. Under condition (3),[29] it is readily verified that there exists an integer $\overline{m}$ such that for all $K > \overline{m}$, player i's best-response to β_i requires Passing (with probability 1) in the first $K-\overline{m}$ moves (at least) (because for some appropriately specified $\overline{m}$, Taking earlier is dominated by never Taking).

Suppose that players 1 and 2 Pass with probability 1 in the first node where they must move. The consistency condition implies that the analogy-based expectation of player i, $\beta_i(\alpha_i) = \lambda^i.Pass + (1-\lambda^i).Take$, should satisfy $\lambda^i \geq \frac{1}{2}$.

Together these two observations imply that the mapping

$$\beta = (\beta_1, \beta_2) \underset{\text{Best-response}}{\rightarrow} \sigma = (\sigma_1, \sigma_2) \underset{\text{Consistency}}{\rightarrow} (\beta_1 \langle\sigma\rangle, \beta_2 \langle\sigma\rangle)$$

has a fixed point such that $\lambda^i \geq \frac{1}{2}$ for $i=1,2$. Given the best-response to such analogy-based expectations, we may conclude.

(2) Suppose player i's strategy requires him to Pass with probability 1 in node $N_i^{(K)}$ for $i=1,2$. By the consistency requirement it should be that player i's analogy-based expectation $\beta_i(\alpha_i) = \lambda^i.Pass + (1-\lambda^i).Take$ satisfies $\lambda^i \geq \frac{1}{2}$ for $i=1,2$. The best-response to β_i is to Pass at least in the first $K-\overline{m}$ where he must move. ■

Proposition 4 (1) shows that irrespective of the length $2K$ of the game, there is an equilibrium (possibly in mixed strategies) in which Take occurs in a finite number of periods toward the end of the game.[30] Proposition 4 (2) shows that there cannot be equilibria in which Take occurs in the middle phase of the game (i.e. between period 3 and period $2K-2$).

Comment 1: A prediction of the analogy setup (at least with the coarsest partitioning and restricting attention to equilibria in which Take never occurs in the first two periods) is that, by increasing the length of CP_K, the length of the end phase - in which Take may occur - can never grow above some fixed and bounded value.

Comment 2: It should be noted that the Subgame Perfect Nash Equilibrium outcome is also an analogy-based expectation equilibrium outcome (in which Player 1 Takes in $N_1^{(K)}$ expecting player 2 to Take in α_1). And that there is another equilibrium in mixed strategies

[29] Since all payoffs are integers satisfying (2), the differences $a_t - a_{t+2}$, $b_t - b_{t+2}$ are no smaller than 2.

[30] When condition (4) does not hold, this may involve an equilibrium in mixed strategies.

in which Take may occur in the first two periods (it is such that each player $i = 1, 2$ plays in mixed strategies in $N_i^{(K)}$ and Takes with probability 1 in all other nodes).

We now consider a slight modification of the environment in which the Subgame Perfect Nash Equilibrium is no longer an equilibrium and Take can only occur toward the end of the game. Specifically, assume the players not only play game CP_K, but also another game that is the same as game CP_K except that there are only two moves corresponding to Player 1 passing or not to the middle of the game and Player 2 passing or not from the middle to the end of the game. Formally, let K be an odd number. Consider the game tree Υ such that in stage 0 Nature selects either game CP_K with probability $\nu_{CP} > 0$ or game F with probability $\nu_F > 0$ where game F is described as follows.

Game F has the same two players $i = 1, 2$ as CP_K and two moves. Player 2 moves in the first node denoted by M_2. At node M_2, player 2 must choose an action in $A_2 = \{Pass, Take\}$. If player 2 Takes, the game ends, players 1 and 2' payoffs are a_{2K} and b_{2K}, respectively. If player 2 Passes, the game moves to node M_1 where it is player 1's turn to move. Player 1 must choose an action in $A_1 = \{Pass, Take\}$. If player 1 Takes, this is the end of the game and the payoffs of players 1 and 2 are a_K and b_K, respectively; if he Passes, this is also the end of the game and the payoffs to players 1 and 2 are a_0 and b_0, respectively. We assume that K is larger than 2 so that $a_0 > a_K > a_{2K}$ and $b_0 > b_K > b_{2K}$.

Also, we assume that each player i uses a single class of analogy. That is,

$$\alpha_i = \left\{(j, N_j^{(k)}), 1 \le k \le K\right\} \bigcup \{(j, M_j)\}$$

and we call $(N, \Upsilon, \%_i, An)$ the associated environment.

Proposition 5 *Suppose that conditions (2) and (3) hold. There exists an integer $\overline{m}$ such that for all $K > \overline{m}$,* all *self-confirming analogy-based expectation equilibria (σ, β) of $(N, \Upsilon, \%_i, An)$ are such that player i Passes with probability 1 in M_i and in every $N_i^{(k)}$, $k = K, ...K - \overline{m}$.*

Proof. In game F, whatever their analogy-based expectation, each player i chooses optimally to Pass. This ensures that the analogy-based expectation of player i, $\beta_i(\alpha_i) = \lambda^i.Pass + (1 - \lambda^i).Take$ satisfies $\lambda^i \ge \nu_F > 0$ for $i = 1, 2$. Given condition (3), this ensures that, for K large enough, the best-response in CP_K of each player i is at least to Pass in the first node where he must move, thus ensuring that $\lambda^i > \frac{1}{2}$ for $i = 1, 2$. We may then conclude using the best-response argument in the proof of Proposition 4. ■

In the above analysis of the centipede game CP_K, we assumed that players use the coarsest analogy partitioning. However, the insight that analogy reasoning may lead players to Pass most of the time in long enough CP_K would in general carry over, even when players use more than one analogy class.

Suppose, for example, that each player i considers two classes:

$$\begin{aligned}\alpha_i^{end} &= \left\{(j, N_j^{(k)}) \text{ such that } k \geq \overline{k}\right\} \\ \alpha_i^{main} &= \left\{(j, N_j^{(k)}) \text{ such that } k < \overline{k}\right\}\end{aligned}$$

according to whether the end phase or the main phase of the game is being considered, and call $(N, CP_K, \succsim_i, An)$ the corresponding environment.

Proposition 6 *There exist $\overline{m}$ and an analogy-based-expectation equilibrium of $(N, CP_K, \succsim_i, An)$ such that, for all K, each player i Passes with probability 1 at least in the first $K - \overline{m}$ nodes where he must move.*

Proof. If $\frac{\overline{k}-1}{\overline{k}} b_{2k} + \frac{1}{\overline{k}} b_{2k+1} > b_{2k+2}$ for all $k \leq \overline{k}$, then Player 1 Taking in the last node $N_1^{(1)}$ is an analogy-based expectation equilibrium outcome (this follows from the analysis in Proposition 3).

Otherwise, using the argument in the proof of Proposition 4, it is readily verified that there is $\overline{m}$ such that for, K large enough, there is an equilibrium (σ, β) satisfying (1) $\beta_i(\alpha_i^{main}) = \lambda^{i,main} \cdot Pass + (1 - \lambda^{i,main}) \cdot Take$ with $\lambda^{i,main} \geq \frac{1}{2}$ for $i = 1, 2$, and (2) player i's best-response to β_i is to Pass with probability 1 at least in the first $K - \overline{m}$ moves. ■

3.3 (Finitely) Repeated Prisoner's Dilemma

Consider the Prisoner's Dilemma PD whose matrix payoff is represented as:

	D	C
D	$0, 0$	$1 + g_1, -l_2$
C	$-l_1, 1 + g_2$	$1, 1$

Game PD

with l_i, $g_i > 0$ for $i = 1, 2$, where each player $i = 1, 2$ has to choose simultaneously an action in $\{D, C\}$. We now consider several variants of repeated PD. The first two variants illustrate how analogy reasoning may give rise to (non-trivial) end effects in the finitely repeated PD. The third variant deals with the infinite repetition.

T-repetition: We first consider the T repetition of PD with no discount factor, and we denote by PD_T the corresponding game tree. Nodes in PD_T correspond to histories of length 0 to T specifying the action profiles played in earlier periods (if any). The history of length 0 is denoted by $\emptyset$, and a history h of length $t > 0$ is $(a^{(1)}, ..., a^{(t)})$ where $a^{(k)} = \left(a_1^{(k)}, a_2^{(k)}\right)$ and $a_i^{(k)} \in \{D, C\}$ stands for the action played by player i in period k.

We consider the following analogy partitioning. Player 1 partitions the set of $(2,h)$ into two classes:[31]

$$\begin{aligned} \alpha_1 &= \{(2,h) \mid h \text{ contains no } D \text{ or } h = \emptyset\} \\ \alpha_1' &= \{(2,h) \mid h \text{ contains at least one } D\} \end{aligned}$$

Player 2 is assumed to use the *finest* partitioning of the set of $(1,h)$. We refer to An as the corresponding partitioning, to $N = \{1,2\}$ as the set of players, and to $\%_i$ as player i's preferences over outcomes in PD_T (as induced by the above matrix function and the no discounting assumption).

Proposition 7 *For T sufficiently large, the path $(a^{(k)})_{k=1}^{T}$ with $a^{(k)} = (C,C)$ for all $k \leq T-2$, $a^{(T-1)} = (C,D)$ and $a^{(T)} = (D,D)$ is the equilibrium path of an analogy-based expectation equilibrium of $(N, PD_T, \%_i, An)$.*

Proof. Suppose the proposed path corresponds to the analogy-based expectation equilibrium (σ, β). By consistency, one should have:[32]

$$\begin{aligned} \beta_1(\alpha_1) &= \frac{T-2}{T-1} \cdot C + \frac{1}{T-1} \cdot D \\ \beta_1(\alpha_1') &= D \end{aligned}$$

For T large enough (so that $\frac{T-2}{T-1}$ is sufficiently close to 1), player 1's best-response σ_1 to β_1 is to play D in the last period T or[33] whenever one (or more) D has been played before. And to play C otherwise.[34]

Player 2's best-response σ_2 to σ_1 is to play D whenever one (or more) D has been played before and in the last two periods $T-1$ and T. And to play C otherwise.[35]

The proof is completed by noting that the path generated by (σ_1, σ_2) corresponds to the assumed path. ■

Remark 1*:* It is interesting to observe that in class α_1 (resp. α_1'), Player 1 behaves in the same way in all histories h reached along the equilibrium path and such that $(2,h) \in \alpha_1$ (resp. α_1').

---

[31] History $h = (a^{(1)}, ..., a^{(t)})$ is said to contain at least one D if there exist $i = 1,2$ and $k \leq t$ such that $a_i^{(k)} = D$.

[32] Note that the two analogy classes of player 1 are reached on this path.

[33] Playing D in the last period T is optimal whatever the expectation (it is a dominant strategy).

[34] To see this, consider (the worst case at) period $T-1$ in which player 1 is supposed to play C. If he plays C (planning to play D next), his expected continuation payoff (given his expectation) is close to $1 + (1 + g_1)$ (corresponding to $(C,D),(C,C)$. If he plays D (anticipating (D,D) will occur afterwards) his expected continuation payoff is close to $(1 + g_1) + 0$ (corresponding to $(D,D),(C,D)$, which is smaller.

[35] Player 2 finds it optimal to play D in period $T-1$ because he knows that in any case player 1 will play D in the last period - he is fully rational.

Remark 2: Suppose Player 2 uses the partitioning An_2 of $\{(1,h)\}$ obtained from An_1 by exchanging the roles of players 1 and 2. Then the path $((C,C),...,(C,C),(D,D))$ can be sustained as an analogy-based expectation equilibrium path.

Stochastic T-repetition: We now consider a variant of PD_T with T periods and no discounting, but such that in each period there is a draw by Nature to determine the value of each g_i for the current period (the values of l_i are assumed to remain constant throughout the game). To fix ideas, we assume that the distributions are independent from period to period and accross players, and that in each period, g_i takes value $\underline{g}$ with probability $\underline{\nu}$ and $\overline{g}$ with probability $\overline{\nu}$ where $\overline{\nu}+\underline{\nu}=1$.[36] We denote by $z^{(t)}$ the joint draw of (g_1,g_2) in period t, and we assume that players are risk neutral. We denote by PD_s the associated game tree, and by $\%_i$ player i's preferences.

Nodes in PD_s correspond to histories of length 0 to T specifying the action profiles played in earlier periods (if any) and the draws by Nature in all periods up to (and including) the current period. That is, the history of length 0 is $z^{(1)}$ specifying the draws g_1 and g_2 for the first period. A history h of length $t>0$ is $((a^{(1)},z^{(1)});...;(a^{(t)},z^{(t)});z^{(t+1)})$ where $a^{(k)}=\left(a_1^{(k)},a_2^{(k)}\right)$, $z^{(k)}=\left(g_1^{(k)},g_2^{(k)}\right)$ and $a_i^{(k)}\in\{D,C\}$ stands for the action played by player i in period k while $g_i^{(k)}\in\left\{\underline{g},\overline{g}\right\}$ stands for the period k draw of g_i.

Each player i partitions the set of (j,h) into two classes:[37]

$$\begin{aligned}\alpha_i &= \{(j,h)\mid h \text{ contains no } D\}\\ \alpha_i' &= \{(j,h)\mid h \text{ contains at least one } D\}\end{aligned}$$

We define $u_T=1+(\underline{\nu}\cdot\underline{g}+\overline{\nu}\cdot\overline{g})$, and the sequence $(u_t)_{t<T}$ recursively by[38]

$$u_t=1+(\underline{\nu}\cdot u_{t+1}+\overline{\nu}\cdot\overline{g}).$$

We assume that $u_T<\overline{g}$ and that no x_t in this sequence is equal to $\overline{g}$ (which is satisfied generically). We define $\overline{m}$ as the integer such that $u_{T-\overline{m}+1}<\overline{g}<u_{T-\overline{m}}$. (Note that $\overline{m}$ is no larger than $\overline{g}$, since $u_t-u_{t+1}>1$ as long as $u_{t+1}<\overline{g}$.)

[36] To keep in line with the class of games considered in Section 2, we assume that the draws of both g_1 and g_2 are immediately revealed to both players. However, this is immaterial for the analysis below.

[37] History $h=((a^{(1)},z^{(1)});...;(a^{(t)},z^{(t)});z^{(t+1)})$ is said to contain at least one D if there exist $i=1,2$ and $k\leq t$ such that $a_i^{(k)}=D$. It is said to contain no D otherwise.

[38] u_t stands for the expected payoff of player i at date $t-1$ when no C previously occurred and player i anticipates that (1) he will play D in the next period if $g_i=\overline{g}$ (or if $t=T$) and that (2) player j plays C if no D previously occurred and D otherwise.

Proposition 8 *For T large enough, the following strategy profile is part of an analogy-based expectation equilibrium of $(N, PD_s, \succsim_i, An)$: For each player i, play D if one (or more) D occurred so far; Otherwise, in all periods t, $t < T - \overline{m}$, play C; in all periods t, $T - \overline{m} < t < T$, play C if $g_i^{(t)} = \underline{g}$ and D if $g_i^{(t)} = \overline{g}$; in period T, play D.*

Proof. Given the assumed strategy profile σ and given that $\overline{m}$ is no larger than $\overline{g}$, for T large enough, the analogy-based expectations $\beta_i \langle \sigma \rangle$ that is consistent with σ should satisfy:

$$\begin{aligned} \beta_i \langle \sigma \rangle (\alpha_i) &\approx C \\ \beta_i \langle \sigma \rangle (\alpha'_i) &= D \end{aligned}$$

It can be checked that the best-response to such a $\beta_i \langle \sigma \rangle$ is indeed σ_i.[39] ■

The logic of the equilibrium is as follows. Players rightly perceive that in class α'_i, i.e. if some D was played earlier, only D can be expected next. In class α_i, player j chooses C most of the time except toward reaching the end of the game: player i' expectation is thus close to C in this class. Given such an expectation, player i considers breaking the sequence of C - by playing D - only when the immediate gain g_i from playing D is not too small relative to the loss incurred by triggering a D sequence. This occurs only toward the end of the game (i.e. in the last $\overline{m} + 1$ periods) when the draw of g_i is $\overline{g}$ (and also in the last period irrespective of the realization of $g_i^{(T)}$).

Remark 1*:* The same result as in Proposition 8 carries over if the two classes considered by player i are now such that α_i contains only those (j, h) such that player i (and not necessarily player j) has never played D so far.

Remark 2*:* The above results suggest that analogy reasoning may provide an explanation for the experimental evidence[40] that players initially cooperate in the finitely repeated prisoner's dilemma and that there is an end effect in which players sometimes behave opportunistically (from period $T - \overline{m}$ to period T).

Infinite repetition: We briefly consider the infinitely repeated prisoner's dilemma in which both players have the same discount factor δ, $g_1 = g_2 = g > 0$ and $l_1 = l_2 = l > 0$. The associated game tree is denoted by PD_δ, and player i's preference is denoted by $\succsim_i$.

As a preliminary comment, we first observe that a cooperative outcome may emerge for some analogy partitioning, but not for others. We let α_i denote player i's partitioning of (j, h) where $h = (a^{(1)}, ..., a^{(t)})$ is a t-length history of action profiles.

---

[39] The sequence u_t has been constructed precisely for that purpose.

[40] See Selten-Stoecker (1986), but also McKelvey-Palfrey (1992), Nagel-Tang (1998).

Suppose player i's partitioning An_i bears *only* over the actions played by player j. That is, if h and h' involve the same sequence of actions of player i, (j, h) and (j, h') belong to the same analogy class. Then the only path that can be sustained as an analogy-based expectation equilibrium of $(N, PD_\delta, \%_i, An)$ is the repetition of (D, D).

The point is that due to player i's partitioning, player i will play D irrespective of his expectation, since there is no effect identified by player i of i's own actions on j's behavior. As a result, only the repetition of (D, D) can emerge.

Of course, if each player i partitions the set of (j, h) according to

$$\begin{aligned} \alpha_i &= \{(j,h) \mid h \text{ contains no } D\} \\ \alpha_i' &= \{(j,h) \mid h \text{ contains at least one } D\} \end{aligned}$$

then the trigger strategy profile in which each player i plays C whenever no D was played so far and plays D otherwise is part of an analogy-based expectation equilibrium of $(N, PD_\delta, \%_i, An)$.[41]

In the above examples, it is readily verified that the analogy-based expectation equilibria (σ, β) of $(N, PD_\delta, \%_i, An)$ are such that σ is a Subgame Perfect Nash Equilibrium of PD_δ. We now show that this need not be the case.

Example 2: Assume that $1 - l > \frac{1}{2}(1+g)$. Player 2 uses the finest partitioning of $(1, h)$. Player 1 partitions $(2, h)$ into two classes:

$$\begin{aligned} \alpha_1 &= \left\{(j,h) \mid h \text{ contains no } (D,C) \text{ and no } k \text{ s.t } a_2^{(k)} = a_2^{(k+1)} = D\right\} \\ \alpha_1' &= \{(j,h) \mid (j,h) \notin \alpha_1\} \end{aligned}$$

Define the following strategy profile σ:

1. For player 1: Play D if at least one (D, C) occurred so far or if player 2 played D in two consecutive periods in the past. Play C otherwise.

2. For player 2: Play D if at least one (D, C) occurred so far or if player 2 played D in two consecutive periods in the past or if the last action profile is (C, C). Play C otherwise.

   And define player 1's analogy-based expectation β_1 as[42]: $\beta_1(\alpha_1) = \frac{1}{2}C + \frac{1}{2}D$ and $\beta_1(\alpha_1') = D$.

Claim 2: For δ close enough to 1, (σ, β) is an analogy-based expectation equilibrium of $(N, PD_\delta, \%_i, An)$. Yet, σ is not a Subgame Perfect Nash Equilibrium of PD_δ.

[41] The associated analogy-based expectations are: $\beta_i(\alpha_i) = C$ and $\beta_i(\alpha_i') = D$.

[42] β_2 can be identified with σ_1 as player 2 uses the finest partitioning.

Proof. The path generated by σ (starting from the nul history) is an alternation of (C, C) and (C, D). Hence, the expression for $\beta_1(\alpha_1)$. ($\beta_1(\alpha_1') = D$ is the correct expectation.) It is readily verified that σ_2 is a best-response to σ_1. After an (D, C) occurs or player 2 played D in two consecutive periods, each player i's strategy requires playing D. Note that even player 1' expectation is correct in this case, and the requirements for equilibrium are satisfied. In all other events and if the last action played by player 2 was C (resp. D), player 1 expects (given β_1 and the ensuing strategy σ_1) to get u^C (resp. u^D) by playing C where

$$\begin{aligned} u^D &= \frac{1}{2}(-l+1) + \delta\frac{u^C}{2} \\ u^C &= \frac{1}{2}(-l+1) + \delta\frac{u^C + u^D}{2} \end{aligned}$$

(As δ gets close to 1, $u^D \approx (-l+1)$ and $u^C \approx 2(-l+1)$.) If player 1 plays D instead of C when the last action of player 2 is C (resp. D), player 1 expects to get $\frac{1}{2}(1+g) + \delta\frac{u^D}{2}$ (resp. $\frac{1}{2}(1+g)$). When $1 - l > \frac{1}{2}(1+g)$, player 1 prefers playing C in both cases.

The strategy profile σ is not a SPNE because player 1 should optimally prefer playing D rather than C on the equilibrium path when the last action profile is (C, C). ■

The effect of player 1's treatment by analogy here is to allow to sustain strategy profiles in which player 1 is not punished after every deviation player 1 may consider. (In the above example, player 1 is not punished if he plays D instead of C when player 2 is supposed to play D.) Player 1 conforms to the prescribed strategy because he perceives an average punishment in case of deviation (even if in some events there would be no actual punishment).

3.4 Bargaining and ultimatum games

In this subsection we wish to illustrate the effect of analogy reasoning in situations where the action spaces are large at least in some nodes. A simple example is provided by the following Take-it-or-Leave-it model. 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.[43] 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 $0 < v^{out} < 1$. Call $N = \{1, 2\}$ the set of players, $\succsim_i$ player i's preferences, and TL the game tree associated with the above setup.

[43] The action space of player 1 in this example is continuous (which is not covered by the framework of Section 2). The analysis presented below can be viewed as corresponding to the limit of the finite grid case as the grid becomes finer and finer. (Alternatively, one can easily extend the definitions of consistency and of analogy-based expectation equilibrium for this specific example.)

Standard analysis suggests that player 1 will propose $(1-v^{out}, v^{out})$ 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 in TL 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:[44]

$$\begin{aligned} \alpha_1^{low} &= \{(2, x) \mid \overline{x} < x \leq 1\} \\ \alpha_1^{high} &= \{(2, x) \mid 0 \leq x \leq \overline{x}\} \end{aligned}$$

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

Proposition 9 *(1) When $1-\overline{x} < v^{out}$, any analogy-based expectation equilibrium is such that there is no agreement: player 1 gets 0, player 2 gets v^{out}. (2) When $v^{out} < 1-\overline{x}$, there is a unique analogy-based expectation equilibrium in which player 1 proposes $(\overline{x}, 1-\overline{x})$, which is accepted by player 2.*

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 $r = low, high$. If $\lambda^{high} > 0$ (resp. $\lambda^{low} > 0$), player 1's best-response to β_1 cannot be to offer $(x, 1-x)$ with $x < \overline{x}$ (resp. $x < 1$). (1) When $1-\overline{x} < v^{out}$, neither $(1, 0)$ nor $(\overline{x}, 1-\overline{x})$ are acceptable by player 2. Only a disagreement can occur. (2) When $v^{out} < 1-\overline{x}$, $\lambda^{high} = 1$, $\lambda^{low} = 0$, player 1 proposing $(\overline{x}, 1-\overline{x})$ and player 2 accepting any offer $(x, 1-x)$ with $1-x \geq v^{out}$ gives rise to an analogy-based expectation equilibrium. It is also immediate to check that there is no other analogy-based expectation equilibrium. ■

Comment 1*:* The analysis of Proposition 9 is pretty similar to the one that would arise if player 1 could only propose a partition offer in $\{(\overline{x}, 1-\overline{x}), (1, 0)\}$. Thus the analogy reasoning here has the effect (through the working of the best-response correspondence) of reducing the offers considered by player 1 in equilibrium.

Comment 2*:* Another setup which would give similar insights is one in which the responder would not distinguish within the set of high offers (i.e., offers x such that $x \leq \overline{x}$) nor within the set of low offers (i.e., offers x such that $x > \overline{x}$).[45] However, in a slightly more elaborate model in which player 1 could not pick a deterministic offer, but could only affect

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

[45]See Dow (1991), Meyer (1991) and Rubinstein (1993) for investigations of coarse informational partitionings of this sort.

the distribution of offers received by player 2, then the two approaches would have different behavioral implications.[46]

To give a further illustration of the effect of analogy reasoning in this kind of contexts, consider the following bargaining game (see Rubinstein 1982). There are two players $i = 1, 2$ and a pie of size 1. Players alternate in making offers. Player 1 starts, and makes an offer $(x_1, 1 - x_1)$, $x \in [0, 1]$ to player 2 who may either accept or reject it. If he accepts, players 1 and 2 get x_1 and $1 - x_1$, respectively. If player 2 rejects the offer, player 2 makes an offer $(1 - x_2, x_2)$ to player 1 and so on. Players are assumed to discount future payoffs according to the same discount factor δ. Call $N = \{1, 2\}$ the set of players, $\succsim_i$ player i's preferences, and R the game tree.

Standard analysis suggests that player 1 will propose $(\frac{1}{1+\delta}, \frac{\delta}{1+\delta})$ and that player 2 will accept it immediately. We now show that when players reason by analogy, an agreement need not be reached immediately resulting in severe efficiency losses.

Nodes in which player 1 (resp. 2) must move are of two types: nodes in which player 1 (resp. 2) must make an offer - we refer to such nodes as h_1^{off} (resp. h_2^{off})-, and nodes h_1^{resp} (resp. h_2^{resp}) in which player 1 (resp. 2) must respond to an offer $(1 - x_2, x_2)$ (resp. $(x_1, 1-x_1)$). We assume that each player i uses three classes to categorize the nodes at which player j must move:

$$\begin{aligned}
\alpha_i^{off} &= \left\{(j, h) \mid h = h_j^{off}\right\} \\
\alpha_i^{low} &= \left\{(j, h) \mid h = h_j^{resp} \text{ and } i\text{'s current offer } x_i \in (\overline{x}_i, 1]\right\} \\
\alpha_i^{high} &= \left\{(j, h) \mid h = h_j^{resp} \text{ and } i\text{'s current offer } x_i \in [0, \overline{x}_i]\right\}
\end{aligned}$$

One can show:[47]

Proposition 10 *Suppose that $\overline{x}_1 + \overline{x}_2 > 1$ and let δ be sufficiently close to 1. The following assessment (σ, β) is an analogy-based expectation equilibrium of $(N, R, \succsim_i, An)$. Player i's strategy: Propose $\overline{x}_i$ for himself (and $1 - \overline{x}_i$ for player j), Accept with probability p_i the offer $1 - \overline{x}_j$ of player j, Reject (Accept) any offer strictly below (above) $1 - \overline{x}_i$. Player i's*

[46] For the sake of illustration, suppose player 1 chooses x (as in the main presentation). When player 1 chooses x, player 2 receives the offer $(x, 1 - x)$ with a large probability, but also any offer $(y, 1 - y)$ with $y > x$ with a small probability. We assume the same partitioning (regarding the offer effectively received by player 2) as in the main text and $1 - \overline{x} < v^{out}$. When the responder has a coarse perception, he will reject any offer $(y, 1 - y)$, $y \leq \overline{x}$, whereas in the analogy setup he will accept offers $(y, 1 - y)$ whenever $1 - y > v^{out}$. Hence, the non equivalence: in the analogy setup, player 1 will pick $\overline{x}$, and sometimes with small probability the deal will be accepted; in the coarse perception setup, there will never be any agreement.

[47] The probability p_i is determined so that player j is indifferent between accepting $1 - \overline{x}_i$ today and having his offer of $1 - \overline{x}_j$ accepted with probability p_i tomorrow (and otherwise accepting $1 - \overline{x}_j$ aftertomorrow).

expectation: $\beta_i(\alpha_i^{off})$ *is the offer* $1-\overline{x}_j$, $\beta_i(\alpha_i^{low}) =$*'Rejects'*, $\beta_i(\alpha_i^{high}) = p_j$ *'Accepts'*$+(1-p_i)$ *'Rejects' where* $p_i = \frac{1-\delta^2}{\delta}\frac{1-\overline{x}_j}{\overline{x}_i-\delta(1-\overline{x}_j)}$.

Comment 1*:* In this equilibrium as δ gets close to 1, each player i gets approximately $1-\overline{x}_j$, which results in an efficiency loss of $\overline{x}_1+\overline{x}_2-1$.

Comment 2*:* In the usual Rubinstein (1982)'s argument, if player j's continuation payoff is (approximately) $1-\overline{x}_i$, player i could get $\overline{x}_i-\varepsilon$ by offering $1-\overline{x}_i+\varepsilon$ to player j. In our analogy setup, player j would also accept such an offer with probability 1. However, due to his analogy partitioning, player i perceives that the probability of acceptance of such an offer is only p_i exactly the same as for the offer $1-\overline{x}_i$. Given such an expectation, player i prefers offering $1-\overline{x}_i$ rather than $1-\overline{x}_i+\varepsilon$.[48]

4 Discussion

4.1 Link with other solution concepts

We have already illustrated the differences between the analogy-based expectation equilibrium and the Subgame Perfect Nash Equilibrium in a number of applications (see Section 3). We have also noted the fundamental conceptual distinction between the notion of information set in multi-stage games with incomplete information and the notion of analogy class as developed in Section 2. (They refer to partitioning of different sets, see Section 2.)

Consider now the notion of imperfect recall in which a player need not remember which actions he chose in the past (see Rubinstein 1991, Piccione-Rubinstein 1997, and also Dulleck-Oechssler 1997 for an application to the centipede game).[49] Consider again the centipede game CP_K described in subsection 3.2. But assume that each player $i=1,2$ does *not* know at which node $N_i^{(k)}$, $k=1,...K$ he currently is (whereas players no longer form their expectations by analogy). For K large enough, an equilibrium in this setting is that each player i Passes with probability 1 in his unique memory/information set $I_i = \left\{N_i^{(k)}, k=1,...K\right\}$.[50]

[48] The insight we get here is similar to the one that would arise if each player i could only propose $1-\overline{x}_i$ to player j. There is some parallel with the effect of finite grids in bargaining as analyzed in van Damme-Selten-Winter (1990) (who focus on the multiplicity issue). However, the grid resulting from the analogy partitioning need not be the same for the two players (nor need it be fine, as considered by these authors). Also, the analogy treatment forces here stationarity in the form of the equilibrium, which is not implied by the finite grid treatment.

[49] One can argue that games with imperfect recall fall in the class of games with incomplete information (see the discussion in Piccione-Rubinstein 1997).

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

Imperfect recall explains in this case why players may Pass all the time in the centipede game. However, it does so by assuming that players do not perceive that there is an end (since players are assumed not to know at which node they currently are). In the analogy approach developed in subsection 3.2, players *do* know at which node they currently are. They Pass initially because they do not have an *accurate* expectation about when their opponent will Take (they only have an expectation about the average behavior of their opponent all over the game). Also, players do perceive that there is an end, as they consider Taking toward reaching the end of the game. Thus the two approaches have a very different interpretation, and only the analogy approach captures (in an endogenous way) the end effect in the finite horizon paradoxes.

Can the solution concept defined in Section 2 receive an alternative interpretation?[51] As we have just suggested, the relation to imperfect recall (or incomplete information) is not clear. But, what about a situation in which players would *erroneously* believe that the other players do not have perfect information, and yet all players would have perfect information. Classical game theory does not provide a clue about how to model such situations. But maybe (restricting attention to two player games), one can interpret the analogy classes α_i defined in Section 2 as the fictitious information sets assigned to player j by player i,[52] and propose the corresponding equilibrium (see Definition 4) as the appropriate concept for this case.

Here is a problem though with this interpretation: Suppose that in equilibrium both (j, h) and (j, h') in α_i are reached with positive probability. And that player j does not behave in the same way in h and h'. In principle, player i could *check* ex post (or statistically after many plays of the game) that player j does not behave in the same way in h and h', violating i's premise that h and h' are in the same information set of j. Would player i continue to hold such a belief about player j's information structure in such a case?[53]

Remark: With this interpretation, if one adds the requirement that player j should behave in the same way in all nodes h and h' reached in equilibrium and such that (j, h) and (j, h') belong to the same α_i, then the play must correspond to that of a self-confirming equilibrium (which need not be the case in general with the analogy-based expectation equilibrium).[54]

[51] I would like to thank a referee for suggesting the following interpretation.

[52] With this interpretation, it seems odd to pool together (j, h) and (j', h') with $j \neq j'$ as the general setup of Section 2 allows it, hence the restriction to two-player games.

[53] If player i does not even think of checking whether player j behaves (statistically) in the same way in h and h' , the proposed concept seems to make sense, but the interpretation then gets closer to that of analogy reasoning in that player i is satisfied with a partial expectation about player j's behavior.

[54] The only self-confirming equilibrium in the centipede game CP_K requires player 1 Taking in the first node (and see the analysis in subsection 3.2).

4.2 Two principles on analogy partitioning

No structure was imposed so far on the analogy categorization. Understanding how individuals categorize contingencies to form their expectations is clearly beyond the scope of this paper (it is at the heart of a large body of the ongoing research of cognitive scientists, see Holyoak-Thagard 1995 and Dunbar 2000, for example). As a modest game-theoretic investigation, we now review two principles (for analogy partitionings) that may alternatively be viewed as attempts to refine the concept of analogy-based expectation equilibrium.

Analogy expectation and similar play:

An appealing idea seems to be 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 player i plays the same behavioral strategy at nodes h and h'. We distinguish according to whether this property should be met only for histories reached along the played path or for all histories.[55]

In the finitely repeated prisoner's dilemma PD_T in which player 1 uses two classes of analogy α_1, α_1', (according to whether or not at least one D was played earlier) and player 2 uses the finest partition, we observed (see Remark 1 after Proposition 7) that for all histories h met along the played path, whenever $(2, h) \in \alpha_1$ (resp. α_1'), player 1 plays the same action C (resp. D) at node h. Thus, the property is met for all histories h *reached* along the played path.[56]

The next example is such that the property is met for all histories (whether on or off the equilibrium path), and yet the play differs from that of the Subgame Perfect Nash Equilibrium:

Example 3: 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 $\{L, R\}$. 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

[55] One might argue that a player is more likely to have doubts about his analogy partitioning if the property is violated on the equilibrium path histories.

[56] The requirement is not met though for the off the equilibrium path history $h = (a^{(t)})_{t=1}^{T-1}$ with $a^{(t)} = (C, C)$ for all t, in which player i would play D (and not C as for the other histories in α_1).

player 1 derives a higher payoff when (U, R) is played in game G than when (U, L) is played in game G'. And 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 (this yields player 1 more than (U, L) in G').

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' rather than (U, L) in game G. ■

All analogy classes must be reached:

Another property that may be of interest is that players structure their analogy classes so that each analogy class is reached with positive probability in equilibrium.[57] The next Proposition provides some insight about the effect of this principle in the centipede game CP_K considered in subsection 3.2.

Proposition 11 *Let (σ, β) be an analogy-based expectation equilibrium of $(N, CP_K, \%_i, An)$ where $N = \{1, 2\}$ denotes the set of players, $\%_i$player i's preferences, and An the partitioning profile used by the players. Suppose that for all k, $\frac{1}{2}a_{k-2} + \frac{1}{2}a_{k-1} > a_k$ and $\frac{1}{2}b_{k-2} + \frac{1}{2}b_{k-1} > b_k$. If σ employs only pure strategies, and* 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 subsection 3.2). 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

[57] A possible psychological rationale for this is that players tend to prefer structuring analogy classes so that expectations can be checked on the equilibrium path (without trembling requirement).

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).[58] 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

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

But given this expectation (and given that for all k, $\frac{1}{2}a_{k-2}+ \frac{1}{2}a_{k-1} > a_k$ and $\frac{1}{2}b_{k-2}+ \frac{1}{2}b_{k-1} > b_k$), 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. ■

4.3 Multiplicity and Analogy-Based Expectation Equilibrium

In this subsection we would like to highlight the implication of analogy reasoning on the multiplicity of equilibria. A first observation is that the analogy treatment may sometimes kill the multiplicity of equilibria that would otherwise prevail. For example, in the infinitely repeated prisoner's dilemma PD_δ, we observed that if one player has an analogy partitioning such that his own actions play no role, then the only equilibrium outcome is the repetition of (D, D). Here, the analogy treatment kills the multiplicity because it does not permit enough conditioning of players' expectations.

A second observation is that the analogy treatment may sometimes create a multiplicity of equilibria by permitting some form of conditioning that would not be possible otherwise. For example, in the finitely repeated prisoner's dilemma in which player 1 uses two classes according to whether or not at least one D was played earlier and player 2 uses the finest partitioning, we saw an equilibrium in which both players play C except in the last two periods (see subsection 3.3.). But, there is also an equilibrium for this partitioning in which both players play D in every period. Here, the multiplicity arises because the analogy treatment permits a conditioning of player 1's expectation (upon whether or not at least one D was played earlier) that would not be possible otherwise (due to the logic of backward induction).

A third observation is that the consistency condition (1) implies non-linearities, thus creating a potential for multiple equilibria, even when players use a single class of analogy[59] and there is a unique equilibrium in the setup without analogy. For example, in the centipede game CP_K in which both players use the coarsest partition and condition (4) holds, we saw that there are two pure strategy analogy-based expectation equilibria (see Proposition 3).

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

[59] Hence, no conditioning of the type just mentioned is at work.

4.4 Related literature

There are very few approaches to analogy in economics. These include the axiomatic approaches of Rubinstein (1988) and Gilboa-Schmeidler (1994) about similarity and case-based decision theory, respectively (which derive representation theorems for decision processes satisfying a number of axioms). These also include the automata theory developed for game theory by Rubinstein (1986), and Abreu-Rubinstein (1988) (see also Samuelson 2000 for a recent contribution with an explicit reference to the analogy interpretation). In the automaton setup, two different histories[60] 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'. It should be noted that none of the above approaches considers the treatment of expectations (as opposed to behaviors) by analogy.

I now discuss a few alternative approaches to bounded rationality and how they relate to (differ from) the analogy-based expectation approach. By bounded rationality, I mean here that either players fail to optimize their true payoffs given their expectations (failure of instrumental rationality) or players fail to have correct expectations (failure of cognitive rationality) or players have a wrong perception about the game being played. Clearly, the analogy-based expectation equilibrium approach challenges the cognitive rationality, but not the other two features of rationality.

In his discussion of small words, Savage (1954) suggests the possibility that a decision maker may not be able to envision the whole complexity of the state space when making a decision.[61] In the analogy-based expectation approach too, players form expectations about the behavior of others by simplifying the node space over which other players must move. One can view the present paper as an attempt to incorporate Savage's small word idea into a game theoretic context.[62]

Simon (see, for example, Simon 1955) is clearly one of the leaders in emphasizing the need to incorporate bounded rationality into economics. One of his main concerns is that the process of decision making be manageable in particular with respect to information gathering. The analogy treatment proposed in this paper shares with this view the worry to ensure that learning is more manageable (as a result of a simplification of what players are supposed to

[60] Or two different games G and G' in Samuelson's setup.

[61] In a one-agent problem this may be viewed as either challenging the cognitive rationality assumption or as an erroneous perception of the decision problem. See also McLeod (2000) for recent research along this line.

[62] A key motivation for my concept is to make learning feasible (by simplifying the predictions on which learning is supposed to bear). Such a motivation is not present in Savage (1955) who develops a more introspective approach.

learn). However, Simon (1955) emphasizes other elements of simplification in the decision making like the use of simplified payoff functions (in the form of satisficing).[63]

More recent approaches include the ε-equilibrium of Radner (1986) and the quantal response equilibrium of McKelvey-Palfrey (1992-1995) (see also Chen-Friedman-Thisse 1997). These approaches assume that players fail to optimize exactly (either they optimize up to ε in Radner or they play any strategy with positive probability in McKelvey-Palfrey), but expectations about what other players might do are assumed to be accurate (in every contingency). Thus these approaches challenge the cognitive rationality, but not the other forms of rationality.[64]

Studies of limited foresight in multi-stage games (see Jehiel 1995-1998-2001) also challenge the cognitive rationality assumption in that players are assumed to form predictions only about a truncated future. However, the form of limitation imposed by limited foresight is substantially different from that imposed by analogy reasoning.[65]

In the context of normal form games, Osborne-Rubinstein (1998) model situations in which players behave as if they were not in a strategic environment. Players experiment each possible action once (or several times), and choose the action that delivered the highest payoff. This approach (at least with this interpretation) as well as that of Greenberg (1996) challenges the view that players have a correct perception of the game being played (see also Camerer 1998 for an experimental account of misperceptions of games).

Finally, other approaches to the finite horizon paradoxes include the crazy type approach (Kreps et al. 1982), the ε-equilibrium (Radner), the quantal response equilibrium (McKelvey-Palfrey), the lack of common knowledge of the termination date (Neyman 1986), the justifiability approach (Spiegler 2000).[66] The ε-equilibrium and quantal response equilibrium approaches were already discussed. None of the other approaches rely on a form of bounded rationality as defined above.

[63] Also, Simon does not emphasize the implications of his approach in strategic environments with several decision makers interacting at the same time.

[64] See also Tesfatsion (1984) for another example of this sort.

[65] For example, limited foresight alone is incapable of explaining that players may Pass in the centipede game. In the precursor of this paper (Jehiel 1999), I combined limited foresight and the analogy idea to analyze the centipede game.

[66] In the finitely repeated prisoners' dilemma, Benoit develops a non-equibrium approach in which expectations about the other player's strategy may be arbitrary. Since no constraint of any sort is placed on these expectations, his approach seems hard to generalize.

References

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

[2] Benoit, J. P (1988): 'A non-equilibrium analysis of the finitely-repeated prisoner's dilemma,' *Mathematical Social Sciences* **16**: 281-287.

[3] Camerer, C. (1998): 'Mental representations of games,' mimeo.

[4] van Damme, E. , R. Selten and E. Winter (1990): 'Alternating bid bargaining with a smallest money unit,' *Games and Economic Behavior* **2**: 188-201.

[5] Dow, J. (1991): 'Search decisions with limited memory,' *Review of Economic Studies* **58**: 1-14.

[6] Dunbar K. (2000): 'The analogical paradox: why analogy is so easy in natural settings, yet so difficult in the psychological laboratory,' mimeo McGill University.

[7] Dulleck, U. and J. Oechssler (1997): 'The absent-minded centipede,' *Economics Letters* **55**: 309-315.

[8] Fudenberg D. and D. Levine (1998): *The theory of learning in games*, MIT Press.

[9] Gilboa, I. and D. Schmeidler (1995): 'Case-based decision theory,' *Quarterly Journal of Economics* **110**: 605-639.

[10] Greenberg, J. (1996): 'Towering over Babel: worlds apart but acting together,' mimeo.

[11] Holyoak K. J. and P. Thagard (1995): *Mental leaps: analogy in creative thought*, Cambridge, MA, MIT Press.

[12] Jehiel, P. (1995): 'Limited horizon forecast in repeated alternate games,' *Journal of Economic Theory* **67**: 497-519.

[13] Jehiel, P. (1998): 'Learning to play limited forecast equilibria,' *Games and Economic Behavior* **22**: 274-298.

[14] Jehiel, P. (1999): 'Predicting by analogy and limited foresight in games,' mimeo CERAS and UCL.

[15] Jehiel, P. (2001): 'Limited foresight may force cooperation,' forthcoming *Review of Economic Studies*.

[16] Kreps, D., P. Milgrom, J. Roberts and R. Wilson (1982): 'Rational cooperation in the finitely repeated prisoner's dilemma,' *Journal of Economic Theory* 27: 245-252.

[17] Kreps, D. and R. Wilson (1982): 'Sequential equilibria,' *Econometrica* 50: 863-894.

[18] MacLeod W. B. (2000): 'Cognition and the theory of learning by doing,' mimeo.

[19] Mc Kelvey R. and T. Palfrey (1992): 'An experimental study of the centipede game,' *Econometrica* 60: 803-836.

[20] Mc Kelvey R. and T. Palfrey (1995): 'Quantal response equilibrium for normal form games,' *Games and Economic Behavior* 10: 6-38.

[21] Meyer, M. (1991): 'Learning from coarse information: biased contests and career profiles,' *Review of Economic Studies* 58: 15-41.

[22] Nagel, R. and F. F. Tang (1998): 'Experimental results on the centipede game in normal form: an investigation on learning,' *Journal of Mathematical Psychology* 42: 356-384.

[23] Neyman, A. (1985): 'Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma,' *Economics Letters* 19: 227-229.

[24] Osborne, M. and A. Rubinstein (1998): 'Games with procedurally rational players,' *American Economic Review* 88: 834-847.

[25] Piccione, M. and A. Rubinstein (1997): 'On the interpretation of decision problems with imperfect recall,' *Games and Economic Behavior* 20: 3-24.

[26] Radner, R. (1986): 'Can bounded rationality resolve the prisoner's dilemma?,' in A. Mas-Colell and W. Hildenbrand, eds. *Essays in honor of Gerard Debreu*, 387-399.

[27] Rosenthal R. (1982): 'Games of perfect information, predatory pricing and the chain-store paradox,' *Journal of Economic Theory* 25: 83-96.

[28] Rubinstein, A. (1982): 'Perfect equilibrium in a bargaining model,' *Econometrica* 50: 97-109.

[29] Rubinstein, A. (1986): 'Finite automata play a repeated prisoner's dilemma,' *Journal of Economic Theory* 39: 145-153.

[30] Rubinstein, A. (1988): 'Similarity and decision-making under risk,' *Journal of Economic Theory* 46: 145-153.

[31] Rubinstein, A. (1991): 'Comments on the interpretation of game theory,' *Econometrica* **59**: 909-924.

[32] Rubinstein, A. (1993): 'On price recognition and computational complexity in a monopolistic model,' *Journal of Political Economy* **101**: 473-484.

[33] Rubinstein, A. (1998): *Modeling bounded rationality*, MIT Press.

[34] Samuelson, L. (2000): 'Analogies, adaptation, and anomalies,' forthcoming *Journal of Economic Theory*.

[35] Savage L. J. (1954): *The Foundation of Statistics*, New York John Wiley and Sons, Inc.

[36] Selten, R. (1975): 'Reexamination of the perfectness concept for equilibrium points in extensive games,' *International Journal of Game Theory* **4**: 22-55.

[37] Selten, R. and R. Stoecker (1986): 'End behavior in sequences of finite prisoner's dilemma supergames,' *Journal of Economic Behavior and Organization* **7**: 47-70.

[38] Simon, H. (1955): 'A behavioral model of rational choice,' *Quarterly Journal of Economics* **69**: 99-118.

[39] Spiegler R. (2000): 'Reason-based choice and justifiability in extensive-form games,' mimeo

[40] Tesfatsion L. (1984): 'Games, goals, and bounded rationality,' *Theory and Decision* **17**, 149-175.

[41] Tversky, A and D. Kahneman (1981): 'The framing of decisions and the psychology of choice,' *Science* **211**: 453-458.