Epistemic Game Theory

Lecture 14

Eric Pacuit

University of Maryland, College Park
pacuit.org
epacuit@umd.edu

May 12, 2014

A. Colman. *Cooperation, psychological game theory, and limitations of rationality in social interaction*. Behavioral and Brain Sciences, 26, pgs. 139 - 198, 2003.

"Rationality has a clear interpretation in individual decision making, but it does not transfer comfortably to interactive decisions, because interactive decision makers cannot maximize expected utility without strong assumptions about how the other participant(s) will behave. In game theory, common knowledge and rationality assumptions have therefore been introduced, but under these assumptions, rationality does not appear to be characteristic of social interaction in general." (pg. 152, Colman)

"Game theorists of the strict school believe that their prescriptions for rational play in games can be deduced, in principle, from one-person rationality considerations without the need to invent collective rationality criteria — provided that sufficient information is assumed to be common knowledge." (p. 142, Binmore)

"Game theorists of the strict school believe that their prescriptions for rational play in games can be deduced, in principle, from one-person rationality considerations without the need to invent collective rationality criteria — provided that sufficient information is assumed to be common knowledge." (p. 142, Binmore)

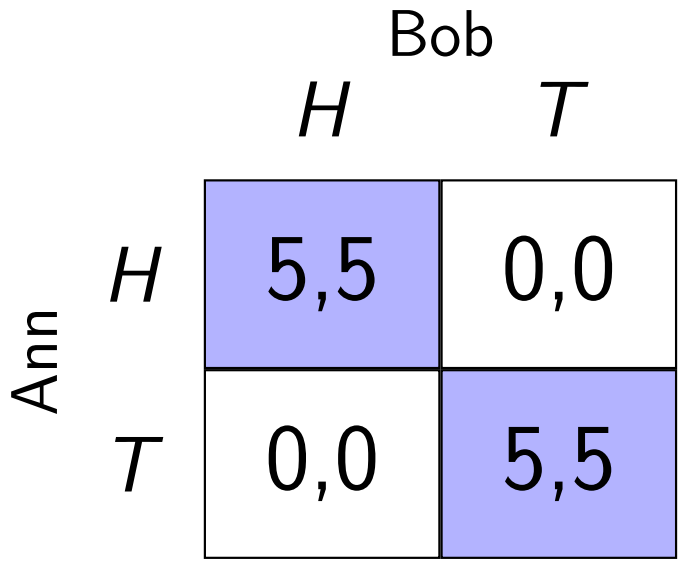
"full rationality is not such a bad assumption; it is a sort of idealization, like the ideas of perfect gas or frictionless motion;...no less valid than any other scientific idealization" (p. 139, Aumann).

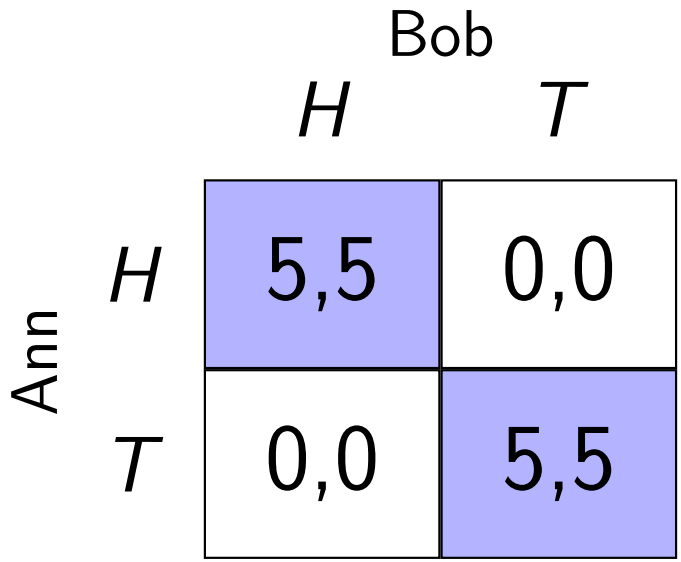
"Von Neumann and Morgenstern (1944)... set out to derive a theory of rational play in games from one of rational individual decision-making. Their successors have not deviated from the faith that this can be done" (pp. 3-4, Hurley and Bacharach).

The following assumptions are fairly standard and are often called common knowledge and rationality (CKR):

- CKR1: The specification of the game, including the players' strategy sets and payoff functions, is common knowledge in the game, together with everything that can be deduced logically from it and from CKR2.
- CKR2: The players are rational in the sense of expected utility (EU) theory, hence they always choose strategies that maximize their individual expected utilities, relative to their knowledge and beliefs at the time of acting. (By CKR1 this too is common knowledge in the game.)

Any rational deduction about the game must (by CKR1) be common knowledge—Bacharach named this the *transparency of reason*. It implies that, if it is uniquely rational for Player 1 to choose Strategy X and Player 2 strategy Y, then X and Y must be best replies to each other, because each player anticipates the other?s strategy and necessarily chooses a best reply to it. Because X and Y are best replies to each other, they constitute an equilibrium point by definition. Therefore, if a game has a uniquely rational solution, then it must be an equilibrium point.


Bob
H
T
Ann
H
T
5,5
0,0
0,0
5,5



In the non-cooperative game of Heads or Tails, rational players are forced to choose arbitrarily, with a probability of successful coordination of $1/2$ and an expected payoff of 2.5. Can they do better than that?

Going beyond the mathematical properties of the game and delving into its psychology, if both players perceive heads to be more salient than tails, in other words if they both recognize (H, H) as a focal point, and if both believe this to be common knowledge, then both will unhesitatingly choose heads, and they will coordinate successfully.

"The salient focal points are obvious in Heads or Tails, the unspecified meeting time, and Hume's problem of the three wines.

"The salient focal points are obvious in Heads or Tails, the unspecified meeting time, and Hume's problem of the three wines. Nevertheless, it turns out that their selection cannot be justified rationally.

"The salient focal points are obvious in Heads or Tails, the unspecified meeting time, and Hume's problem of the three wines. Nevertheless, it turns out that their selection cannot be justified rationally. Gilbert (1989b) showed that "if human beings are — happily — guided by salience, it appears that this is not a consequence of their rationality" (p. 61) and that "mere salience is not enough to provide rational agents with a reason for action (though it would obviously be nice, from the point of view of rational agency, if it did)" (p. 69, emphasis in original)."

The focal point of Heads or Tails is obviously (Heads, Heads), and to clarify the argument, let us assume that the players have previously agreed on this, so it is common knowledge.

The focal point of Heads or Tails is obviously (Heads, Heads), and to clarify the argument, let us assume that the players have previously agreed on this, so it is common knowledge.

Under the CKR2 rationality assumption, Player I will choose heads, given any reason for believing that Player II will choose heads, to ensure a payoff of 5 rather than 0.

The focal point of Heads or Tails is obviously (Heads, Heads), and to clarify the argument, let us assume that the players have previously agreed on this, so it is common knowledge.

Under the CKR2 rationality assumption, Player I will choose heads, given any reason for believing that Player II will choose heads, to ensure a payoff of 5 rather than 0.

But in the absence of any reason to expect Player II to choose heads, Player I has no reason to choose it or not to choose it.

The fact that (Heads, Heads) is a focal point is not a valid reason for Player I to choose heads, because heads is best only if Player II chooses it also. Because the salience of (Heads, Heads) does not give Player I a reason to choose heads, it cannot give Player I a reason to expect Player II to choose heads.

The fact that (Heads, Heads) is a focal point is not a valid reason for Player I to choose heads, because heads is best only if Player II chooses it also. Because the salience of (Heads, Heads) does not give Player I a reason to choose heads, it cannot give Player I a reason to expect Player II to choose heads.

Both players are in exactly the same quandary, lacking any reason for choosing heads in the absence of a reason to expect the co-player to choose it. The argument goes round in circles without providing the players with any rational justification for playing their parts in the focal-point equilibrium, in spite of its salience and intuitive appeal.

This is an excellent example of the fundamental thesis of this article, that the concept of utility maximization cannot be applied straightforwardly to interactive decisions.

Team Reasoning

A team-reasoning player maximizes the objective function of the set of players by identifying a profile of strategies that maximizes their joint or collective payoff, and then, if the maximizing profile is unique, playing the individual strategy that forms a component of it.

Team Reasoning

A team-reasoning player maximizes the objective function of the set of players by identifying a profile of strategies that maximizes their joint or collective payoff, and then, if the maximizing profile is unique, playing the individual strategy that forms a component of it.

"Team reasoning is simple and intuitively compelling but profoundly subversive of orthodox decision theory and game theory, both of which rest on a bedrock of *methodological individualism*.....Team reasoning is inherently non-individualistic and cannot be derived from transformational models of social value orientation."

Stackelberg Reasoning

"Suppose my co-player could read my mind, or at least anticipate my strategy choice. If I chose H, my co-player would best-reply H, and I'd receive my optimal payoff of 6. If I chose L, then my co-player would best-reply L, and I'd receive 3. In this version of the game, I'd maximize my payoff by choosing H. My co-player knows all this, because we share a common faculty of reason, so in the actual game, I should choose H and expect my co-player to do likewise."

Both team reasoning and Stackelberg reasoning may help to explain the payoff-dominance phenomenon. In addition, Stackelberg reasoning can be shown to predict focal-point selection in pure coordination games in general (Colman 1997). Team reasoning may even offer a partial explanation for cooperation in social dilemmas. Any alternative explanations of these phenomena would have to invoke other nonstandard psychological game-theoretic processes that have yet to be discovered.

Chain-store paradox: A chain-store has branches in 20 cities, in each of which there is a local competitor hoping to sell the same goods. These potential challengers decide one by one whether to enter the market in their home cities. Whenever one of them enters the market, the chain-store responds either with aggressive predatory pricing, causing both stores to lose money, or cooperatively, sharing the profits 50-50 with the challenger.

Intuitively, the chain-store seems to have a reason to respond aggressively to early challengers in order to deter later ones. But Selten's (1978) backward induction argument shows that deterrence is futile.

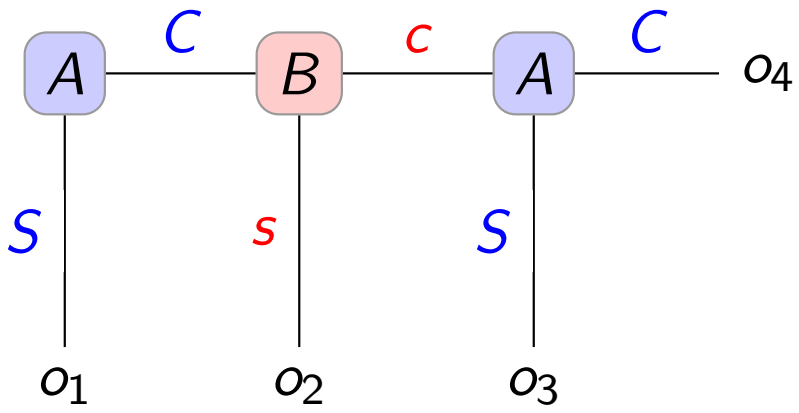
"I would be very surprised if it failed to work. From my discussions with friends and colleagues, I get the impression that most people share this inclination. In fact, up to now I met nobody who said that he would behave according to [backward] induction theory. My experience suggests that mathematically trained persons recognize the logical validity of the induction argument, but they refuse to accept it as a guide to practical behavior." (Selten 1978, pp. 132 - 33)

A. Rubinstein. *Comments on the Interpretation of Game Theory*. Econometrica, 59:4, pgs. 909 - 924, 1991.

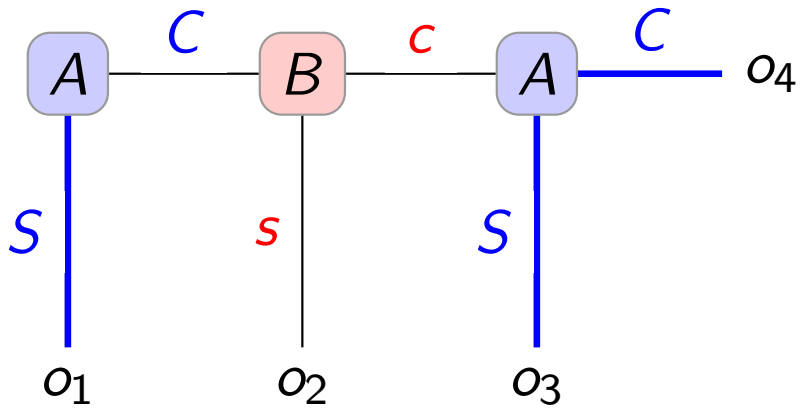
"I approach this paper with the view that game theory is not simply a matter of abstract mathematics but concerns *the real world*. This does not mean that the object of game theory is to predict behavior in the same sense as the sciences do, or indeed, that it is capable of such a function. I view game theory as an analysis of the concepts used in social reasoning when dealing with situations of conflict. It is an abstract inquiry into the function and logic of social institutions and patterns of behavior. " (Rubinstein, pg. 909)

"I aim to endorse the view that equilibrium strategy describes a player's plan of action, as well as those considerations which support the optimality of his plan (i.e., preconceived ideas concerning the other players' plans) rather than being merely a description of a "plan of action". (Rubinstein, pg. 910)

Strategies in Sequential Games

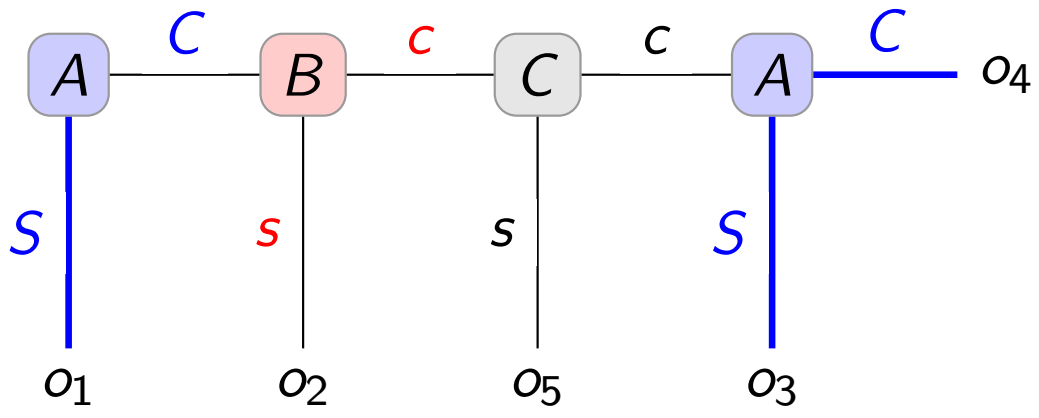


Strategies in Sequential Games



A strategy encompasses not only the player's plan but also his opponents' beliefs in the event that he does not follow that plan.

Strategies in Sequential Games



Players B and C hold the same belief about player A's future behavior, not only on the equilibrium path, but also after player A has deviated.

Mixed Strategies

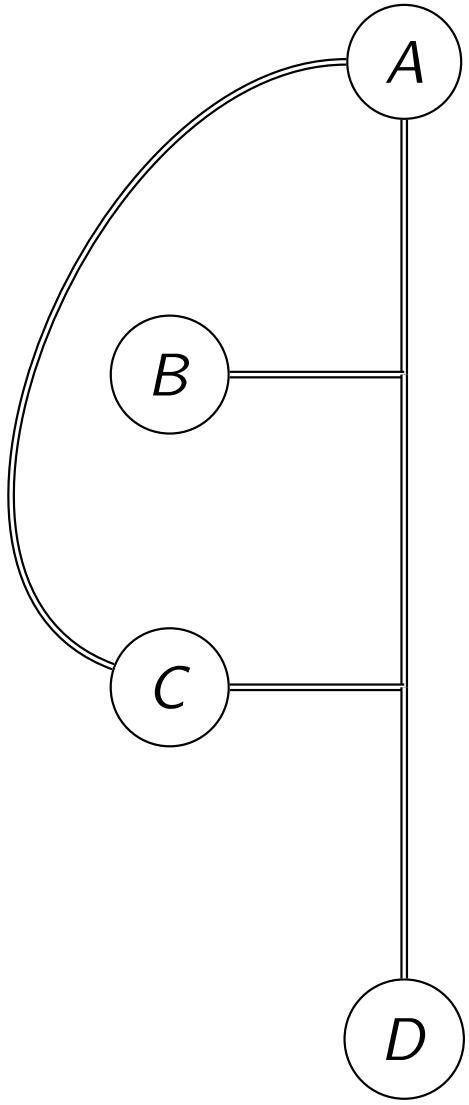
"We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason four each action we take. Outside of Las Vegas we do not spin roulettes."

- One can think about a game as an interaction between large populations...a mixed strategy is viewed as the distribution of the pure choices in the population.
- Harsanyi's purification theorem: A player's mixed strategy is thought of as a plan of action which is dependent on private information which is not specified in the mode. Although the player's behavior appears to be random, it is actually deterministic.
- Mixed strategies are beliefs held by all *other* players concerning a player's actions.

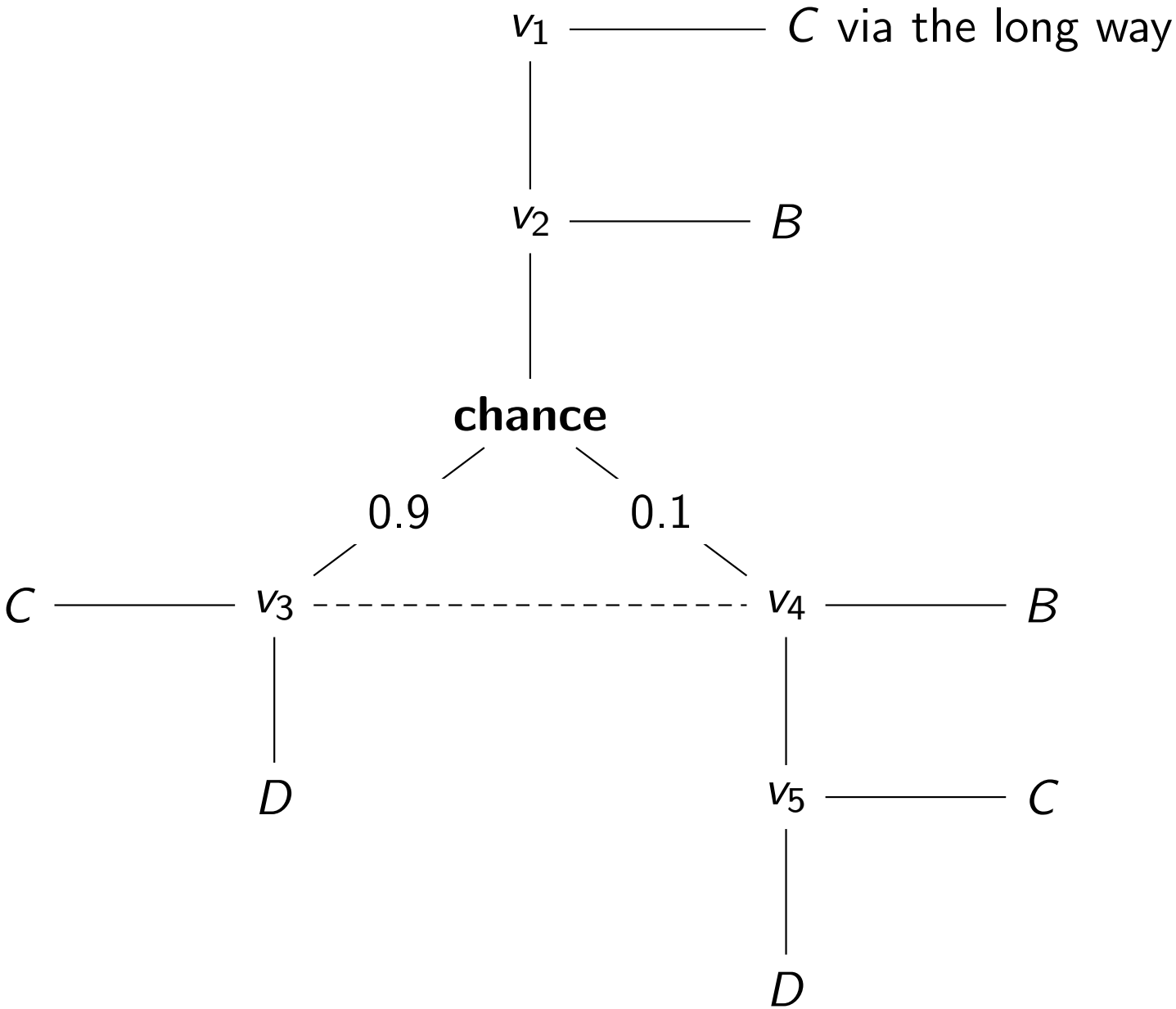
Games with Limited Memory

A drive is at A and wished to reach C. He could drive to C by the long route, which would bring him to his destination directly without having to make any further decisions, or he could use a short but unmarked road, in which case he would have to make a turn at the second intersection. If he arrives at B or if he misses the second intersection and reaches D he will be stuck in traffic jams and hence waste several hours returning to C.

The driver knows that he is able to identify the turn to B but that when he arrives at the turn C, he will become confused and belief that there is a 10% chance that he has not yet passed the first intersection.



A driver needs to make at most 3 decisions.


v_1
C via the long way
v_2
B
chance
0.9
0.1
C
v_3
v_4
B
D
v_5
C
D

Can a strategy in this game be interpreted as a plan of action? No....In the above extensive form game, there is a path (v_1, v_2, v_4, v_5) in which the driver has to make 4 decisions....The decision at v_5 is not part of a plan of action made at A. It is added to the game form merely to allow us to discuss the player's reasoning in the state of doubt.

Strategies are *not* Plans of Action

- In sequential games, such an interpretation does not apply to the part of a player's strategy which supposedly describes a player's planned actions should he deviate from his original plan.
- A mixed strategy can rarely be interpreted as a set of instructions
- In games with potential loss of memory, the game theoretic strategy has to be interpreted as including hypothetical plans which can never be realized.

Games vs. Game Situations

A **game** is a description of strategic interaction that includes

- actions the players *can* take
- description of the players' interests (i.e., preferences),
- description of the "structure" of the decision problem

Games vs. Game Situations

"We adhere to the classical point of view that the game under consideration fully describes the real situation — that any (pre) commitment possibilities, any repetitive aspect, any probabilities of error, or any possibility of jointly observing some random event, have already been modeled in the game tree." (pg. 1005)

E. Kohlberg and J.-F. Mertens. *On the strategic stability of equilibria*. Econometrica, 54, pgs. 1003 - 1038, 1986.

Games vs. Game Situations

Notice that it is rare that a situation involving a conflict of interests is described clearly and objectively by a set of rules. The exceptions I can think of are "games" in the colloquial sense. Unless the game instructions appear on the box bought at "toys 'r' us", I cannot see how we can avoid the interpretation of a game form as an abstract summary of the players' actual perceptions of the complicated situations they are in. "

(Rubinstein, pg. 917)

Games vs. Game Situations

" Formally, a game is defined by its strategy sets and payoff functions. But in real life, many other parameters are relevant; there is a lot more going on. Situations that substantively are vastly different may nevertheless correspond to precisely the same strategic game. For example, in a parliamentary democracy with three parties, the winning coalitions are the same whether the parties each hold a third of the seats in parliament, or, say, 49 percent, 39 percent, and 12 percent, respectively. But the political situations are quite different. The difference lies in *the attitudes of the players, in their expectations about each other, in custom, and in history*, though the rules of the game do not distinguish between the two situations. (pg. 72, my emphasis)

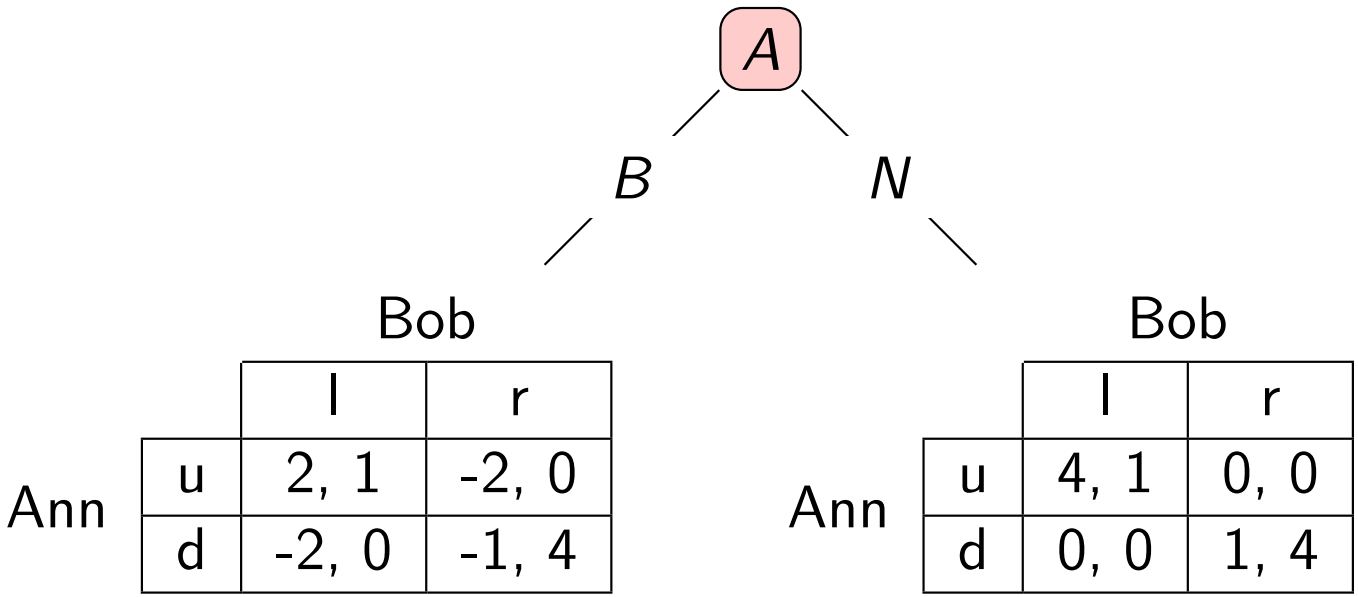
R. Aumann, Robert and J. Dreze. *Rational expectations in games*. American Economic Review, 98(1): 72 86, 2008..

Example: Infinite Horizon Games

By using infinite horizon games we do not assume that the real world is infinite. Models are not supposed to be isomorphic with reality. An infinitely repeated game is meant to assist in analyzing situations where players examine a long-term situation without assuaging a specific statues to the end of the world. In contrast, the finitely related game model corresponds to a situation in which the finite period enters explicitly into the player's considerations.

"If we adopt the view that a game is not a rigid description of the physical rules of the world, then a game-theoretic model should include only those factors which are perceived *by the players* to be *relevant*. Modelling requires intuition, common sense, and empirical data in order to determine the relevant factors entering into the players' strategic considerations and should thus be included in the model. This requirement makes the application of game theory more an art than a mechanical algorithm." (Rubinstein, pg. 919)

Example: Burning Money



Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

Example: Burning Money

		Bob			
		ll	lr	rl	rr
Ann	Bu	2, 1	2, 1	-2, 0	-2, 0
	Bd	-2, 0	-2, 0	-1, 4	-1, 4
	Nu	4, 1	0, 0	4, 1	0, 0
	Nd	0, 0	1, 4	0, 0	1, 4

"If disposing of the dollar were a relevant consideration in the players' perception of the situation, then the result would (probably) make sense. However, I cannot believe that any reasonable person would consider a pre-game disposal of a dollar to be relevant in the analysis of the battle of the sexes. It is my opinion that a formal description of the situation should exclude the choice of disposal even in cases where a description of the game is given by a referee who specifies the possibility of disposing of the dollar (recall that there is rarely a referee)"

Burning the Money is "Irrelevant"

- The disposal decision does not affect the payoffs of the players in the BoS
- The disposal decision does not reveal any unknown information...a sensible conclusion might be that a player who throws the dollar out of the window is just "crazy"
- The disposal decision is not a part of a game which is identical or even similar to the choice problem which Ann has to confront int he BoS. Thus, whether or not Ann disposes the money cannot provide information from which Bob learns about Ann's behavior in the BoS.

Cheap Talk

"If instead of being able to throw a dollar out of the window, Ann is allowed to throw a bill worth nothing out of the winder (or nod her head), or even state "I am going to play u", then the process of successive elimination of weakly dominated strategies is not powerful and all equilibria of the battle of the sexes would survive. It is my impression that although language plays a crucial role in resolving conflicts, game theory has so far been unable to capture this role."

Regularities

Are one-shot games isolated events?

Regularities

Are one-shot games isolated events?

Do players take into account the effect of their choice today on similar future games in which they will participate? If they ignore the effect of their behavior on future games, then the framework of the game-theoretic *one-shot* game is appropriate. If the players calculate the effect of their behavior on future games, the *repeated* game framework is appropriate. If the players use their past experience to speculate about other players' future behavior without taking into account the effect of their own behavior, then we are dealing with *dynamics*.

Regularities

What happens when equilibria do not exist?

Regularities

What happens when equilibria do not exist?

If what we are trying to model in game theory are situations in which we expect regular behavior, then it is not true that all descriptions of the world should have an equilibrium....the nonexistence of a solution concept in pure strategy does not necessarily mean that we should look for stochastic explanations.

Regularities

Observed regularities depend on the language employed.

There exists a widespread myth in game theory, that it is possible to achieve a miraculous prediction regarding the outcome of interaction among human beings using only data on the order of events, combined with a description of the players' preferences over the feasible outcomes of the situation.

There exists a widespread myth in game theory, that it is possible to achieve a miraculous prediction regarding the outcome of interaction among human beings using only data on the order of events, combined with a description of the players' preferences over the feasible outcomes of the situation.

Deductive arguments cannot by themselves be used to discover truths about the world. Missing are data describing the processes of reasoning adopted by the players when they analyze a game. Thus, if a game in the formal sense has any coherent interpretation, it has to be understood to include explicit data on the players' reasoning processes. Alternatively, we should add more detail to the description of these reasoning procedures. We are attracted to game theory because it deals with the mind. Incorporating psychological elements which distinguish our minds from machines will make game theory even more exciting and certainly more meaningful.

Broader Perspectives

F. Dietrich and C. List. *Mentalism versus behaviourism in economics: a philosophy-of-science perspective*. manuscript, 2012.

I. Gilboa, A. Postlewaite, L. Samuelson and D. Schmeidler. *Economic models as analogies*. The Economic Journal, 2014.

Economic theory seeks to explain the social and economic behaviour of human (and sometimes other) agents. It usually does so by (i) ascribing, at least in an 'as if' mode, certain mental states, such as beliefs and desires, to the agents in question and (ii) showing that, under the assumption that those agents are rational, the ascribed mental states lead us to predict, and thereby to 'rationalize', the behaviour to be explained.

Are the ascribed mental states (e.g., subjective probability and utility functions)

1. mere re-descriptions of behavioral patterns and perhaps instrumentally useful constructs for organizing and making sense of empirical regularities (behaviorism)
2. representations of real mental/psychological phenomena, no less existent in the world than the (also not directly observable) electrons, neutrinos, and electromagnetic fields postulated in the natural sciences? (mentalism)

Applying revealed preference theory to game theory

D. Hausman. *Revealed Preference, Belief, and Game Theory*. Economics and Philosophy, 16:1, pgs. 99-115, 2000.

A. Lehtinen. *The Revealed-Preference Interpretation of Payoffs in Game Theory*. Homo Oeconomicus, 28:3, pgs. 265 - 296, 2011.

This can't be right...

"Modern utility theory makes tautology of the fact that action B will be chosen rather than A when the former yields a higher payoff by *defining* the payoff of B to be larger than the payoff of A if B is chosen when A is available." (Binmore, pg. 169)

K. Binmore. *Game Theory and the Social Contract: Playing Fair*. The MIT Press, 1994.

Reading the Normal Form

		Bob	
		L	R
Ann	U	1,1	3,0
	D	0,2	2,3

The numbers must represent the subjective preferences, not the revealed preferences.

1. The only evidence that should be used to test economic theories is evidence about peoples choice behaviour.
2. The content of any economic theory consists solely in its choice-behavioural implications; two theories that are choice-behaviourally equivalent should be seen as equivalent simpliciter.
3. Any economic theory should take the form of a representation of choice behaviour, and that representation should ideally take the form of attributing to the agents the maximization of some objective function.

Gul and Pesendorfer. *The case for mindless economics*. 2008.

Four Misconceptions

1. the misconception of a fixed evidence base,
2. the evidence/content conflation
3. the "unobservable, therefore non-existent" fallacy
4. the maximization dogma

A fixed evidence base

"In short, the idea that the evidence base of a particular scientific discipline should be fixed once and for all lacks any justification, given the history of science and the experience of other scientific disciplines.

A fixed evidence base

"In short, the idea that the evidence base of a particular scientific discipline should be fixed once and for all lacks any justification, given the history of science and the experience of other scientific disciplines. Rather, the evidence base of any science is changeable and dynamic, and there is no reason why economics should be an exception.

A fixed evidence base

"In short, the idea that the evidence base of a particular scientific discipline should be fixed once and for all lacks any justification, given the history of science and the experience of other scientific disciplines. Rather, the evidence base of any science is changeable and dynamic, and there is no reason why economics should be an exception. Accordingly, even if there was a period in the history of economics when peoples choice behaviour was the only evidence used to test theories, there is no principled reason why other kinds of evidence — from people's verbal reports and communicative behaviour to physiological and nueroscientific evidence — could not also be relevant." (pg. 7)

Evidence/content conflation

"But even if the *evidence base* of economic theories were restricted to observable choice behaviour alone — and, as we have seen, there is no principled reason why it should be —

Evidence/content conflation

"But even if the *evidence base* of economic theories were restricted to observable choice behaviour alone — and, as we have seen, there is no principled reason why it should be — it would not follow that the content of any economic theory should consist solely in its choice-behavioural implications. Rather, the *content* of a theory can, and often does, go well beyond its evidence base." (pg. 8)

Unobservable therefore non-existent

"But even if we were to suspend our criticism of the assumption that only choice behaviour is observable in economics, it should be obvious, as a matter of logic, that, from the fact that a particular entity or phenomenon is not observable, it does not follow that this entity or phenomenon does not exist.

Unobservable therefore non-existent

"But even if we were to suspend our criticism of the assumption that only choice behaviour is observable in economics, it should be obvious, as a matter of logic, that, from the fact that a particular entity or phenomenon is not observable, it does not follow that this entity or phenomenon does not exist. And the conclusion that the entity or phenomenon does not exist follows even less from the fact that something is not currently observable. Sometimes we can have strong indirect evidence for something, even though it is not directly observable." (pg. 9)

The maximization dogma

"Which *form* of a theory best explains human behaviour is a contingent, empirical question, which can be settled only by actual scientific research, not by methodological stipulation.

The maximization dogma

"Which *form* of a theory best explains human behaviour is a contingent, empirical question, which can be settled only by actual scientific research, not by methodological stipulation. Just as it has turned out to be wrong — given Einsteins general theory of relativity — that space and time must necessarily be Euclidean (as Immanuel Kant, for example, assumed), so there is no *a priori* reason to think that the explanation of social and economic behaviour must necessarily be based on the maximization of a single objective function. For example, an empirically adequate theory might model agents as being governed by a more complex system of constraints." (pg. 10)

1. Epistemological 'revealed preference' thesis: Our body of evidence for a theory in economics — the set of observation sentences — is restricted to agents choice behaviour.
2. Ontological 'revealed preference' thesis: The ontological commitments of any theory in economics — or at least those ontological commitments that we are entitled to take seriously — are restricted to agents choices and choice dispositions and therefore exclude mental states.

Premise 1: Some mental states, such as beliefs and preferences, are technically among the ontological commitments of our current best theories of economic decision making

Premise 1: Some mental states, such as beliefs and preferences, are technically among the ontological commitments of our current best theories of economic decision making

Premise 2: In any normal science, the criterion for whether a theoretically postulated entity, property, or relation is to be treated as corresponding to a real entity, property, or relation in the world is whether it is among the ontological commitments of our current best theory or theories in the relevant area (assuming we have no special reasons to doubt those theories themselves)

Premise 1: Some mental states, such as beliefs and preferences, are technically among the ontological commitments of our current best theories of economic decision making

Premise 2: In any normal science, the criterion for whether a theoretically postulated entity, property, or relation is to be treated as corresponding to a real entity, property, or relation in the world is whether it is among the ontological commitments of our current best theory or theories in the relevant area (assuming we have no special reasons to doubt those theories themselves)

Premise 3: Economics is a normal science

Premise 1: Some mental states, such as beliefs and preferences, are technically among the ontological commitments of our current best theories of economic decision making

Premise 2: In any normal science, the criterion for whether a theoretically postulated entity, property, or relation is to be treated as corresponding to a real entity, property, or relation in the world is whether it is among the ontological commitments of our current best theory or theories in the relevant area (assuming we have no special reasons to doubt those theories themselves)

Premise 3: Economics is a normal science

Conclusion: The mental states that our best economic theories ascribe to economic agents are to be treated as corresponding to real phenomena (unless we have special reasons to doubt those theories themselves).

"Consequently, the only way to avoid the mentalistic conclusion would be to insist on having special doubts about our economic theories themselves, despite their status as our current best scientific theories in the relevant area. But those asserting such doubts would then have to explain what evidence underpins them." (pg. 19)

"One might think that the difference between mentalism and behaviourism is a purely metaphysical matter, which is of little significance for the practice of economics itself.

"One might think that the difference between mentalism and behaviourism is a purely metaphysical matter, which is of little significance for the practice of economics itself. But this impression is misleading. That the difference matters also in practice can be seen by revisiting the empirical underdetermination problem, the problem that there can exist two or more distinct theories that are empirically equivalent but logically incompatible."

The polite dinner-party guest: Given a choice between a large, a medium-sized, and a small apple, a dinner-party guest (who at home would choose larger apples over smaller ones) chooses the medium-sized apple (for politeness). If the large apple is no longer available while the medium-sized and small ones are, the guest chooses the small apple (again for politeness).

"Does this mean that there is no fact of the matter as to what the correct explanation is? Both our psychological understanding and the practices of other cognitive and behavioral science suggest that there e*can* be a real difference between different rival explanations, despite their choice-behavioural equivalence." (pg. 20)

Models of decision making

- A *paramorphic model* 'describes the empirical phenomena of interest correctly, but the processes underlying the empirical phenomena are not matched by processes in the model'.
- A *homeomorphic model*, by contrast, has the property that 'not only its empirical phenomena match reality, but also its underlying processes do so'.

Models of decision making

- A *paramorphic model* 'describes the empirical phenomena of interest correctly, but the processes underlying the empirical phenomena are not matched by processes in the model'.
- A *homeomorphic model*, by contrast, has the property that 'not only its empirical phenomena match reality, but also its underlying processes do so'.

The behaviouristic account of preferences (and other mental states such as beliefs) is often criticized for what it fails to deliver:

1. it fails to say anything about human psychology and motivation, from which it is explicitly disconnected;
2. it fails to provide adequate foundations for normative economics, as it gives at most an impoverished account of human well-being, says nothing about fundamental desires and needs, and renders interpersonal comparisons of utility impossible (all of which may matter for policy-making); and
3. it fails to 'explain' behaviour in a non-circular way, since behaviour is 'explained' by preferences (or other attributes) that are in turn defined in terms of behaviour.

"...our critique should convince also those who view economics as a science of choice behaviour alone, devoid of any further psychological or normative goals. Our naturalistic argument shows that even if one is not interested in mental states as such, one's theory of choice may well have to take them on board. A theory *of choice* may have to be a theory *about more than choice*." (pg. 25)

I. Gilboa, A. Postlewaite, L. Samuelson and D. Schmeidler. *Economic models as analogies*. The Economic Journal, 2014.

- The assumptions of economics (and game theory) are false
- Scientific fields can sometimes be reduced, at least in principle, to another. One typically finds heavier reliance on mathematics as one moves down the reduction chain. " Economics engages in mathematical analysis that appears in general to be more sophisticated than that employed by psychology or even biology.
- Psychologists are careful to define the scope of applicability of their models very precisely. By contrast, economists tend to find their models useful in a wide variety of examples, viewing the latter as special cases of their models.

Case-based vs. Rule-based reasoning.

Case-based vs. Rule-based reasoning.

"We suggest that economic reasoning is partly case-based, and that one role of theory is to enrich the set of cases. That is, the analysis of a theoretical model can be viewed as an "observation" of a new case. Such a case is not real, but is a *gedankenexperiment*, an observation that is arrived at by pure logic. An observation of this type is new only to the extent that one has not thought about it before." (pg. 8)

E.g., Akerlof's "lemons market" paper.

E.g., Akerlof's "lemons market" paper.

"Despite the fact that this example can be stated as a mathematical result, it may be more useful to think about it as a case rather than as a general rule. As stated, the example can be viewed as the claim, "I have observed a case in which idealized agents, maximizing expected utility, with the following utility functions and the following information structure, behaved in such and such a way".

E.g., Akerlof's "lemons market" paper.

"Despite the fact that this example can be stated as a mathematical result, it may be more useful to think about it as a case rather than as a general rule. As stated, the example can be viewed as the claim, "I have observed a case in which idealized agents, maximizing expected utility, with the following utility functions and the following information structure, behaved in such and such a way". The relevance of this observation for prediction will depend on the perceived similarity between the idealized agents and the real agents one is concerned with, the similarity between the situation of the former and that of the latter, and so forth. " (pg. 8)

E.g., the Ultimatum Game

E.g., the Ultimatum Game

"If one conceives of the model as a general rule, one would have to conclude that the rule was violated, and perhaps re-define its scope of applicability. By contrast, if the theoretical analysis is construed as a case, as is the experimental result, the two coexist peacefully.

E.g., the Ultimatum Game

"If one conceives of the model as a general rule, one would have to conclude that the rule was violated, and perhaps re-define its scope of applicability. By contrast, if the theoretical analysis is construed as a case, as is the experimental result, the two coexist peacefully. Given a new prediction problem, an economist who is asked to make a prediction would have to ask herself, 'is this real problem more similar to the theoretical analysis, assuming common knowledge of rationality with purely monetary payoffs, or is it more similar to the experiment?' " (pg. 10)

"...we argue that the field values axiomatic derivations because axiomatizations and, more generally, equivalence theorems, can be powerful rhetorical tools. The standard view of science leaves little room for rhetoric: theories are confronted with the data, and should be tested for accuracy. By contrast, case-based view of science lets rhetoric occupy center stage: scientists only offer cases, and these should be brought to bear upon prediction problems, where similarity and relevance should be debated as in a court of law. With this openly-rhetorical view of science, the importance of axiomatizations is hardly a mystery."

"One way to facilitate the task of finding second-order analogies is to use a standard language.

"One way to facilitate the task of finding second-order analogies is to use a standard language. One may view a "paradigm" or a "conceptual framework" as consisting of a language that is supposed to be able to describe a large set of cases, coupled with certain principles for prediction.

"One way to facilitate the task of finding second-order analogies is to use a standard language. One may view a "paradigm" or a "conceptual framework" as consisting of a language that is supposed to be able to describe a large set of cases, coupled with certain principles for prediction. For example, the game-theoretical paradigm in economics starts with the language of players, strategies, information sets, outcomes, beliefs, and utilities. This language is somewhat abstract, but it allows economists to see cross-contextual analogies more easily....one replaces terms such as "voters", "buyers", "candidates", and "sellers" with the more abstract "players".

"One way to facilitate the task of finding second-order analogies is to use a standard language. One may view a "paradigm" or a "conceptual framework" as consisting of a language that is supposed to be able to describe a large set of cases, coupled with certain principles for prediction. For example, the game-theoretical paradigm in economics starts with the language of players, strategies, information sets, outcomes, beliefs, and utilities. This language is somewhat abstract, but it allows economists to see cross-contextual analogies more easily....one replaces terms such as "voters", "buyers", "candidates", and "sellers" with the more abstract "players"....In other words, a standard language allows one to see more similarities without resorting to second-order analogies. A paradigm may thus be useful even if it produces no rules."

"one gives up the claim to formulate a general theory, so that accuracy is not an issue, but aims to have a language that describes a wide range of phenomena and allows for higher order analogies.

"one gives up the claim to formulate a general theory, so that accuracy is not an issue, but aims to have a language that describes a wide range of phenomena and allows for higher order analogies. Thus, rule-based reasoning is discarded in favor of case-based reasoning, and, in return, the latter becomes very powerful.

"one gives up the claim to formulate a general theory, so that accuracy is not an issue, but aims to have a language that describes a wide range of phenomena and allows for higher order analogies. Thus, rule-based reasoning is discarded in favor of case-based reasoning, and, in return, the latter becomes very powerful. The claim we are trying to make is that this is the direction taken by much of microeconomic theory in the past few decades, using game theory as the standard model, and generating insightful analogies rather than accurate rules." (pg. 31)