Games and models for games
DGL, 2015
Robert Stalnaker

A game is a sequence of interacting decision problems, normally with more than one agent. The problems interact in that the ultimate outcome (the payoffs for the different agents) will normally depend jointly on the actions of the different players.

An *extensive form* representation of a game is a finite tree with the (non-terminal) nodes representing choice points, labeled by the player who has the choice at that point. The terminal nodes are labeled by a sequence of utility values for the different players, representing the payoff for that player in the case where the game concludes at that point.

In some cases, one of the "players" is *Nature*, whose choices are determined by chance. When a node is labeled with Nature as the player, the branches from the node are labeled with probability values, summing to one.

In some cases, different nodes for a single player are linked together in *information sets*, representing a situation in which the player does not know at which of the nodes she is located. Each node in an information set will have the same number of branches, and the branches will be linked.

A game Γ in *strategic form* is a structure $\langle N, \langle C_i, u_\mathrm{i}\rangle_{i \in N}\rangle$ where N is the set of players, C_i is the set of strategies available to player i, and u_{ii} is player ii's utility function, taking strategy profiles (members of $C = \times_{i \in N} C_i$) to real numbers. The sets, N and each C_i, are finite.

A *model* for a game is a representation of a particular playing of the game. It includes information about the beliefs and degrees of belief of each player at each point in the game, and about the choices that are made, at each point, by each player (and the choices that they would have made at points that are not reached in that particular playing of the game.) A model may also include a representation of the belief revision policies of each of the players: how they would revise their beliefs and degrees of belief if they were to be surprised by the actions of one of the other players.

Here is one way to define a model:

A model M for a game Γ is a structure $\langle W, \mathbf{a}, \langle R_i, P_i, S_i\rangle_{i \in N}\rangle$, where W is a state space, or a set of possible worlds, including all of the possibilities that are compatible with the beliefs of any of the players; **a** is a member of W, the actual world of this particular playing of the game; R_i is a binary doxastic accessibility relation for player i: xR_iy says that possible world y is compatible with what player i believes in possible world x; P_i is a normalized additive measure function that determines player i's partial beliefs in each possible world in the following way: the degree to which i believes proposition ϕ in world x, $P_{i,x}(\phi)$, is defined as the ratio of $P_i(\phi \cap \{y{:}xR_iy\})$ to $P_i(\{y{:}xR_iy\})$; Si is a function taking possible worlds to strategy choices for player ii: $S_i(x)$ is the strategy that player i plays in possible world x. *Propositions* (or to use the statistician's terminology, *events*) are subsets of W.

The resources of a game model allow us to define a rich range of propositions: that player *i* is rational, that player *i* believes (with probability 1) that player *j* is rational, that it is common belief among the players that they all are rational, etc. One can then define the class of models in which a given proposition is true in the actual world of the model. For example, we can define the class of models in which it is common belief that all players are rational.

A *solution concept* is a class of strategies or of strategy profiles that (according to the concept) are the recommended or required (under certain specified conditions) for rational players of the game in question. A model-theoretic definition of a solution concept is given by defining a class of models for the game – models that satisfy the specified conditions. The strategies that count as solutions are those that are played in some model meeting the specified conditions. For example, a strategy is *rationalizable* if it is played in a model in which there is common belief that all players are rational (where rationality is defined as maximizing expected utility).

Model-theoretic definitions of a solution concept correspond to algorithmic definitions in the way that model-theoretic definitions of satisfiability for a logic correspond to proof-theoretic definitions of consistency.

Roger Myerson on solution concepts: "There are a variety of questions that we can use to evaluate these various refinements. Does the proposed solution concept satisfy a general existence theorem? Is the set of solutions invariant when we transform the game in a way that seems irrelevant? Does the solution concept exclude all equilibria that seem intuitively unreasonable? Does it include all the equilibria that seem intuitively reasonable? Does the intuitive logic of the solution concept seem compelling as a characterization of rational behavior?" (Myerson, 240-41.)

Some reading:

Aumann, R. (1976) 'Agreeing to disagree', *Annals of Statistics*, **4**: 1236-1239.

Aumann, R. (1999) 'Interactive epistemology', *Journal of Game Theory*, **28**: 263-314.

Battigalli, P. and G. Bonanno (1999) 'Recent results on belief, knowledge and the epistemic foundations of game theory, *Research in Economics*, **53**: 149-225.

Myerson, R. (1991) *Game Theory: the Analysis of Conflict*. Harvard University Press.

Stalnaker, R. (1996) 'Knowledge, belief and counterfactual reasoning in games', *Economics and Philosophy,* **12**: 133-163.

Stalnaker, R. (1997) 'On the evaluation of solution concepts', in M. Bacharach *et al*, *Epistemic Logic and the Theory of Games and Decisions*. Kluwer Academic Publishers, 345-64.