Game Equivalence

Note for van Benthem Workshop

Alistair Isaac

May 4, 2009

1 Background on Game Equivalence

The most famous work on game equivalence is Thomson (1952) "Equivalence of Games in Extensive Form." Two more recent works which serve as inspiration for the present project are Bonanno (1992) "Set-Theoretic Equivalence of Extensive-Form Games" and van Benthem (2001) "Extensive Games as Process Models."

Thomson frames the debate by considering a large space of games (extensive form games[1]) and a well defined property which can be assigned to each member of that set (the corresponding strategic form game[2]). Thomson then demonstrates first that the space of extensive form games can be partitioned

[1]For those who are unfamiliar with them, extensive form games are basically trees, decorated with

1. an assignment of a single player to each node (the player whose turn to act it is at that point in the game)
2. equivalence relations between nodes labeled with the same player (that player's "information set" at that stage in the game)
3. a function from end nodes and players to payoffs of some form

[2]Again, for those who are unfamiliar, a strategic form game simply lists the sets of outcomes a player can force to obtain, usually in matrix form. The key difference, then, between extensive form and strategic form games is that strategic form games suppress any temporal information. Figures 2 and 3 below offer examples of games in extensive form, and figure 4 gives the corresponding strategic form.

into equivalence classes, where members of a class share their corresponding strategic forms, and second that a set of four transformations are adequate to map any member of a given class into any other member.

Thomson's approach to equivalence has a long history in mathematics and physics. In a nutshell, the two essential conceptual points are these: (1) two formal objects can only be said to be equivalent *with respect to some specified property*; (2) the relationship between this specified property and the formal objects is illuminated by providing the set of transformations under which the specified property remains *invariant*.

Two points of interest immediately arise from these considerations. First, that equivalence is a *relative* notion: formal objects can only be described as equivalent *relative to* some specified property or set of transformations. Second, that this notion of equivalence is purely *descriptive* and *conditional*: *if* we consider this property, *then* these two games are equivalent; no implications are made about which properties are more important than others for determining equivalence.

I emphasize these two points as much of the game theory literature seems to be confused about the relationship between the two. The attitude taken here is that one *first* decides how to model a scenario by picking an appropriate game form; one *second* examines the normative question of which solution concept is appropriate for that game form. Game theorists sometimes argue, however, that two games *should* be treated as equivalent because this equivalence will constrain the space of solution concepts. This discussion often goes hand in hand with the view that there is some absolute (rather than merely relative) notion of equivalence. Here is a passage from Bonanno (1992) exemplifying this type of reasoning:

> . . . when is it that two extensive games are "essentially the same",

> in the sense that rational players would make the "same" choices in the two games? It seems that the only satisfactory way of answering this question is to start from an extensive-form solution concept (or, even better, an explicit definition of rationality) and then define two extensive games to be strategically equivalent if and only if they have the same solution(s). (p. 431)

The attitude taken here is different. The decision of how to formally model an informal real world scenario (an intuitive "game") is a *pragmatic* one. This pragmatic decision must be made *first* before a debate can be had about the appropriate application of the norms of rationality in solving the game. So, from the present standpoint, it would be conceptually confused to allow solution concepts, or analyses of the notion of rationality, to influence one's notion of game equivalence and, correspondingly, one's pragmatic decision about how to model a scenario. In fact, it is my view that an approach such as that suggested by Bonanno begs the question: it assumes an analysis of rationality before a formal framework can be specified in which the question of what rationality amounts to can even be asked.

2 The Project

The goal of the present project is to extend Thomson's result in two directions. First, by considering a larger space of games than Thomson considers. Second, by considering a number of different properties (as opposed to just corresponding strategic form) for partitioning this larger space of games. Following Thomson, however, we will take as relevant properties the corresponding instance of some less detailed game form. Hopefully this will become clear in the sequel.

Since the project is purely descriptive, we make no value judgments about which properties are interesting. In fact, we will let this question be answered pragmatically, by those suggestions which have already been made in the literature. However, in the interests of keeping the project realistic, we will restrict our considerations to game forms which treat both outcomes and moves as *discrete*.[3] The game forms under consideration at the present stage of the project are these:

1. Extensive Form + absentmindedness
2. Extensive Form
3. Set Theoretic Form
4. Strategic Form
5. Reduced (Normal) Strategic Form

Extensive form and strategic form should be familiar (and were defined very loosely in footnotes 1 and 2). Here are some remarks on the remaining forms:

Absentmindedness occurs when two nodes in the same information set fall on the same history.[4] We can interpret absentmindedness as involving a situation in which the agent loses track of where she falls in time. This can be contrasted with *imperfect recall*. Imperfect recall occurs if the agent forgets what move she made. Absentmindedness occurs if the agent forgets whether or not she has made a move. We distinguish extensive form games which allow absentmindedness from those which do not as the equivalence work by Thomson and Bonanno depends upon a definition of extensive form games which prohibits absentmindedness. Also, Tomohiro and I have already done some work (within the context of epistemic temporal logic) on equivalence relations between trees which exhibit absentmindedness and those which do not.

[3]This rules out, for example, bargaining games, which often allow a continuous space of possible outcomes.

[4]A "history" is just a path through the tree.

Reduced strategic form is just strategic form with all redundant moves removed. Strictly speaking, Thomson's results hold for reduced strategic form rather than strategic form as some of his transformations introduce redundant moves.

Set theoretic form is discussed by Bonanno and van Benthem. Essentially it is a generalization of strategic form which includes the temporal structure of the game, but does not include all structure captured by full blown extensive form. The essential insight is this: when a player makes a move, she is choosing between sets of outcomes. All the player can "see" of the game from this standpoint is a sequence of such choices.

One way of thinking about the relationship between extensive form and set theoretic form is this: extensive form prioritizes the causal structure of a game (captured by the tree). This causal structure is then labeled with the agents' information states. Set theoretic form prioritizes the epistemic structure of the game, the set of possibilities agents perceive as available to them at each move.

[*Questions for workshop participants: is this an interesting set of game forms to consider? Are there obvious candidates we have left out, but should be included (e.g. Bayesian games)? Are there candidates we have included but are totally uninteresting (e.g. non-reduced strategic games)?*]

So, what we desire, then, is a space of games which is "maximally expressive" in the sense that for each member of the space, a unique instance of each of the five game forms described above can be assigned. If we can devise such a space of games, then we will have induced five partitions on this space. This would allow us to move to the next stage of the project, discovering the transformations under which these five properties remain invariant. Here, we can appeal to the results of Thomson and Bonanno. Thomson suggests four transformations over extensive form games under which the corresponding normal form remains invariant. Bonanno shows that a generalization of one of Thomson's transformations is adequate to characterize the invariance of set theoretic form over the space of extensive form games. In my work with Tomohiro, we characterized a transformation which takes extensive form games exhibiting absentmindedness into those which do not. Work needs to be done on finding a transformation which moves in the other direction, however.

The rest of this document deals with the problem of identifying a space of game forms which is "maximally expressive" in this sense.

3 "Maximally Expressive" Games

In order to define a general game form, we need an idea of the kind of information included in game representations. Usually, game forms involve a sequence of distinct time steps at each of which a decision can be made by some player. Each (non-trivial) decision restricts the set of possible outcomes. Outcomes also are treated as distinct entities, individuated by some function for each player. Usually, this is a function into **Re** and is interpreted as that player's payoff should the outcome obtain. However, a variety of different functions could play an intermediary role between the outcome itself and the payoffs assigned to agents, so long as these functions provide information about the outcomes which can distinguish them for decision making purposes. For example, one could distinguish outcomes by a function into possible worlds or sets of propositions. The function could also be relational, inducing a partial order on the set of outcomes. In the sequel, we use a simple payoff function for the sake of simplicity.

Agents in general do not have perfect informational access to facts about the game. In particular, an agent may be ignorant of the moves of another player, or that agent may have a defective memory, and thus become ignorant of where they fall in the temporal structure of the game. Acknowledging the possibility of imperfect players demands that we distinguish two questions a game form should answer for each agent, at each point in time:

1. *What possibilities can the agent realize?*
2. *What possibilities does the agent* think *she can realize?*

The first question demands that a game form model the *causal structure* of a scenario; the second demands that it model the *epistemic structure*. The causal structure of a game is most naturally modeled via a tree. A variety of different strategies have been used to model epistemic structure. One popular approach is to label the tree with indistinguishability relations. Another approach distinguishes epistemic states by identifying them with sets of possibilia - the idea is that an agent's decision amounts to picking amongst a set of outcomes (e.g. van Benthem and Bonnano). We can use this latter mechanism to a different purpose, however.

Essentially, at each node in the causal tree, some agent is presented with the choice to act. If we use the suggestion of Bonnano and Benthem, we can characterize this *experience* by a function from nodes into players and collections of subsets of outcomes. The player must choose one of these sets, i.e. she must restrict the possible outcomes of the game in order to have performed an action. Essentially, at each step in the causal structure, a player is faced with a strategic form-like decision. These decision problems may or may not match with the actual choice being made in the causal structure of the game. If a decision problem does not match, then the agent does not correctly perceive the effect of her actions at that decision point, i.e. exhibits some form of *bounded rationality*.

A general causal+epistemic game is a tuple

$$\Gamma = \left\langle N, H, \{\pi_i\}_{i\in N}, \{\equiv_i\}_{i\in N}, \{A_i\}_{i\in N}, \{E_i\}_{i\in N} \right\rangle$$

where

1. N is the set of players

2. H is a set of sequences (histories) which form a tree, this means

   (a) $\emptyset \in H$

   (b) if $(a^k)_{k=1,\dots,n} \in H$ and $m < n$, then $(a^k)_{k=1,\dots,m} \in H$

   The a history $(a^k)_{k=1,\dots,n} \in H$ is *terminal* if there is no a^{n+1} such that $(a^k)_{k=1,\dots,n+1} \in H$. Call the set of all terminal histories Z and the set of all non-terminal histories $H \setminus Z$.

3. $\pi_i : Z \to \mathbf{Re}$ is the payoff function for each player $i \in N$

4. $\equiv_i$ is an equivalence relation on Z such that for each $x, y \in Z$, $x \equiv_i y$ *iff* $\pi_i(x) = \pi_i(y)$ for all $i \in N$. This gives us a notion of identity between outcomes with the same payoffs.[5]

5. A_i is a family of partitioned subsets of $Z/\equiv_i$. Each subset represents the set of outcomes player i considers possible at some step in the game and the partition represents the possible actions which player i perceives as available to her to reduce the remaining possibilia.

6. $E_i : j \times H \setminus Z \to A_j$ for all $i, j \in N$. F_i captures player i's beliefs about the perceptions of each player at each stage in the causal structure of the game.[6]

[5] Technically speaking, this allows for the possibility that agents differentiate outcomes only with respect to their own payoffs. This is commensurate with an interpretation which merely orders outcomes for each player. Alternately, perhaps, we may wish to add more structure, providing a single equivalence relation, one which holds only if all agents agree on the payoffs assigned to two outcomes. This would allow for strategies by which players take the outcomes of opponents into account. Although the choice here is important for solution concepts, it is irrelevant for the following discussion—all that matters is that we have a notion of equivalence between outcomes.

[6] Too much expressive power? Maybe we only need the agent's own experiences, not her assignment of experiences to others. Consider how this power is used to define simultaneity below, though.

[*Questions for workshop participants: is this notion too strong or too weak? Should it be supplemented with some constraints on how the experience functions interact with the causal structure? Unless the two are reasonably closely connected, no reasonable decisions can be made on the basis of experience, i.e. experience becomes irrelevant if it does not interact with the causal structure of the game in an informative manner. This introduces the very interesting question of when an imperfect agent because* so *imperfect that no analysis of what they should (rationally) do can be made sense of.*]

For each of the five game forms listed in section 2, we would like:

1. to associate a unique instance of that game with each possible causal+epistemic form.

2. to provide the set of transformations over causal+epistemic forms such that the corresponding game form remains invariant.

First, however, let's examine two examples to test the expressive power of the causal+epistemic form. In both cases, we focus only on non-obvious aspects of the model.

4 Example of Absentmindedness

The "Absent Minded Driver" (introduced in Piccione and Rubinstein (1997) "On the Interpretation of Decision Problems with Imperfect Recall") describes a motorist who can't distinguish between two intersections on his way home due to inebriation. The traditional model is given in Fig. 1. The indistinguishability of choice points x and y is captured by placing the two nodes within the same information set. In causal+epistemic form, we designate the only player, 1, construct the tree exactly as here, but capture player 1's uncertainty via the

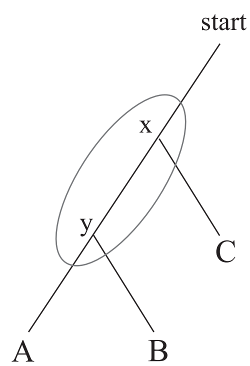

Figure 1: The Absent Minded Driver

experience function E_1. In particular,[7]

$$E_1(x) = E_1(y) = \{\{A\}, \{B, C\}\}$$

The idea here is just a) player 1 has the same experience at nodes x and y, and b) that the content of that experience is that he can either go straight, eventually ensuring outcome A, or he can turn, bringing about outcome B or outcome C.

5 Example of Simultaneity

Bonanno (1992) discusses two extensive form games which he thinks, properly speaking, should not be distinguished. In the first game, player one moves first, then, player two, ignorant of player one's move, makes a move. In the second game, the order of play is reversed, but the causal path to outcomes is manipulated so that player one, moving second, chooses between the exact same two possibilities as in the first game.

[7] Strictly speaking, the function needs a player as argument as well, i.e. $E_1(1)(x) = \{\{A\}, \{B, C\}\}$. We suppress reference to the player here as there is only one player in the game.

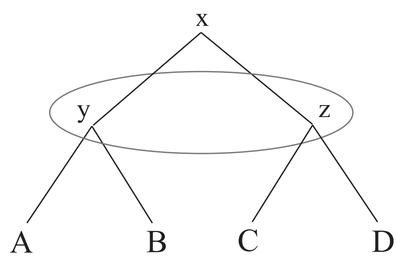

Figure 2: "player 1 moves first"

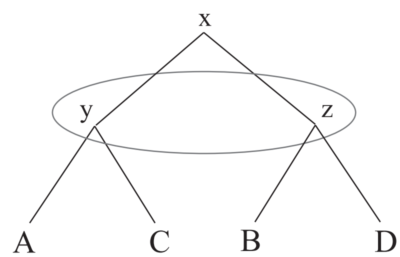

Figure 3: "player 2 moves first"

Bonanno's point is a normative one. Because the ordering of moves is indistinguishable to the players, it should not factor into any solution concept. Furthermore, he points out that there may be scenarios in which moves actually are simultaneous. In these circumstances, it seems undesirable to introduce an arbitrary ordering, especially if there is a danger it will affect the solution concept.

	player 2	
player 1	A	B
	C	D

Figure 4: strategic normal form for both games

Our endeavor is descriptive, however. It would be ideal if we could distinguish all three scenarios: player 1 goes first; player 2 goes first; players 1 and 2 play simultaneously. The E_i functions allow us to distinguish these scenarios in an intuitive fashion.

Remember that $E_i(j)(x) = A$ means that player i models (= believes?) player j as experiencing the choice scenario defined by A. The special case where $i = j$ defines a player's own experience at that node on the tree. If $E_i(j)(x) = \emptyset$ then player i does not model player j as experiencing a decision point at node x.

Consider the causal structure in figure 2. If the nodes represent steps in time, and if these steps are so close that player 1 and player 2's moves are effectively simultaneous, then the respective experience of each player should extend over the entire interval. In other words, $E_1(1)(x) = E_1(1)(y) = E_1(1)(z) = \{\{A, B\}, \{C, D\}\}$ and $E_2(2)(x) = E_2(2)(y) = E_2(2)(z) = \{\{A, C\}, \{B, D\}\}$. If instead, however, the experience of the respective players is separated by a distinct gap, if, in other words, one experiences playing before the other experiences playing, then $E_1(1)(x) = \{\{A, B\}, \{C, D\}\}$, but $E_1(1)(y) = E_1(1)(z) = \emptyset$. Likewise, $E_2(2)(x) = \emptyset$ and $E_2(2)(y) = E_2(2)(z) = \{\{A, C\}, \{B, D\}\}$. (Notice, also, that there are many more distinctions that could be drawn, i.e. one player could believe / experience the moves as simultaneous, while the other experiences them as occurring at distinct temporal steps.

In fact, we may wish to define any two consecutive nodes on a tree as occurring simultaneously *iff* the experience of all players involved is identical at both nodes.

$$h\bar{S}h' \iff h' = he \wedge \forall i, j \in N \left[E_i(j)(h) = E_i(j)(h')\right]$$

6 Future Directions

Although the epistemic possibilities are endless in the above defined causal+epistemic form, only a small subset of them make much sense - basically, this is an unsurprising consequence of the fact that the more imperfect the modeled agent, the less sense can be made of applying the notion of rationality to her strategies. Furthermore, the notion of a "maximally expressive" game defined above is really only of interest if it can be used as a basis for the described equivalence project. Here are some questions for future discussion / investigation:

1. Is this the right way for us to be approaching game equivalence?

2. If not, what should we be doing instead?

3. Is causal+epistemic form too general - i.e. does it include too much extra modeling information for our purposes?

4. If so, how can it be simplified / made more efficient?

5. If not, how much needs to be said about which constraints need to be satisfied to rule out the more outlandish possibilities? Do we even need to rule these possibilities out?

6. Are there more game forms we should be considering than those on the list above?

7. Once we've decided on the list of game forms relevant to what we're doing and the appropriate "maximally expressive" form, then we can move on to the equivalence issue:

    (a) How can we assign a unique member of each game form to each "maximally expressive" form? (Hopefully, we can define the "maximally expressive" form such that this is essentially obvious.)

    (b) Under what transformations does this assignment stay invariant?