Draft 30/4/2016

Could the backward induction controversy be a metaphorical problem?

Ramzi Mabsout
Assistant professor & chair
Economics department
American University of Beirut

Abstract

The backward induction controversy in game theory emerged and practically ended in the 1990s. The protagonists however did not converge to an agreement for explaining the controversy. Why is this the case if opposing sides had access to the same modeling techniques and empirical facts? In this paper I offer an explanation to this controversy and to its unsettled end. It is argued that the answer is not to be found in the modeling claims made by the opposing protagonists but in the tacit metaphors they operate under. Metaphor in Greek means the transfer of meaning from one domain to another. Aristotle defines metaphor as giving a "thing a name that belongs to something else" (Poetica 1457b). The meaning of metaphors has not changed much since then in contrast to models which are comparatively new and still not well understood modes of scientific reasoning. Metaphors provide an independent standard of evaluation from models. The aim of this paper is to utilize game theory to provide a case study that illustrates how modeling, by itself, is not sufficient to overcome the controversies it generates. Metaphor, as a distinct mode of reasoning, could however successfully explain them. The controversy of backward induction in game theory in particular can be used as testing ground for this hypothesis. The paper frames the controversy in terms of metaphor choice to provide a common qualitative reasoning framework to those involved in it. This results in the identification of three different domains—mathematical logic, game theory, and the economic world--each connected to the other via different metaphors. The controversy on backward induction is placed in, and tentatively explained by, this framework.

JEL: B4, B16, C72

1. Introduction

How we think about such questions depends on the predilections we bring to an inquiry, on our suppositions about what will count as an answer, on our explanatory preferences. Keller (2002)

If metaphors are as important as their advocates argue they are their function in science is for cognitive purposes and not merely ornamental. As a cognitive concept metaphors should contribute to our knowledge of the world (Boyd 1979 [1993]) and function "whenever some phenomenon of cognition is conceptualized or explained through the use of metaphor" (Hoffman, et. al.1990:177). The process by which metaphors work is increasingly well understood in natural[1] and social[2] sciences. In biology, psychology, physics, chemistry, economics, and sociology metaphors it is claimed have helped theory formulation, posit novel hypotheses, construct models, explain phenomena, solve problems, and interpret empirical results.

Metaphors are qualitative and procedural --not exacting logically, "never complete, precise, or literal mappings"[3], nor even fixed in time and across contexts. Even if not currently acknowledged in economics, metaphors are formal conceptual tools too, in the sense that they can be used in both teaching and research. They operate on various levels. First, economists themselves can be under the spell of a meta-metaphor, also known as a root, constitutive, or theme metaphor, which explains the origins of the other metaphors they use; second, the model itself can be seen as a metaphor; third, metaphors can be found within the model connecting its various inputs to bridge the gap between the economic world, theory, and the model.[4] One implication of these views is that metaphors will not vanish once an idea has been mathematically modeled. Instead, metaphors are necessary to the mathematical formalization and its interpretation.

Metaphors' most basic definition is that they create a mapping between two distinct domains. In science metaphors' primary function is establishing links between scientific language and the world (Kuhn 1979 [1993:539]). In economics this mapping connects the mathematical model or syntactic structure to the

---

[1] See Emmeche and Hoffmeyer (1991), Pulaczewska (1999), and Jeppson et. al. (2012).

[2] See Bicchieri (1989), Boumans (1999), Bronk (2008), Brown (1977), Clarke et. al. (2014), Fernandez-Duqe and Jonhson (1999), Hodgson (1995; 2005), Klamer and Leonard (1994), Leary (1990), McCloskey (1983; 1994), and Mirowski (1988; 1989; 1994).

[3] Hodgson (1995).

[4] Assuming that theory precedes the model and the model is an articulation of the theory.

economic world (which, it may be noted, economic theory could have helped create performatively). Without the metaphor, the structure consists of uninterpreted syntactical relationships. There is no need to think further than the supply and demand diagram to illustrate the working of the metaphor. In this (geometric or algebraic) model the metaphor maps the usually linear equation (more specifically its intercept and slope coefficients) to feature of the world (quantities and prices of TVs say). Without that mapping the model is silent about the world. The mapping generated by the metaphor can then permit empirically meaningful explanations, possibly even predictions. This does not imply the model is (reducible to) a metaphor (as McCloskey (1990), Brown (1977), and Bronk (2009) contend). Models are more complex creatures than metaphors and they are not so well understood.

This paper explores a game theoretic modeling controversy on the use of backward induction (BI) to compute game theoretic equilibria. The modeling controversy ended in an unexplained deadlock. The back and forth between participants ends abruptly so to speak as neither side grants grounds or agrees on what was actually the subject of the dispute. The controversy has also spawned a micro industry that feeds on what assumptions and definitions justify BI. I shall have more to say on this towards the end of the paper.

The chosen controversy aims to elucidate (or not) the workings of metaphors in economic modeling disputes. A mode of thinking is needed to analyze this controversy. Use could be made of a model to explain this modeling dispute between economists but this brings about a circularity in reasoning. As Hands (2001:391) notes, "the simple fact is that if one wants to "evaluate" something then one needs a standard of evaluation that is based on something other than the thing being evaluated." In this paper use is made of metaphors since they provide a standard of evaluation different from those prevalent in modeling. To evaluate modeling with modeling generates reflexivity and the loss of independent standards of evaluation. An additional factor that militates for the use of metaphors over models in explaining and evaluating the controversy is that the latter are not exactly defined. Metaphor however have had the same consistent definition since Aristotle namely giving a "thing a name that belongs to something else" (Poetica 1457b) . Morgan and Morrison (1999) remark that "there remain a significant lacuna in the understanding of exactly how models in fact function to give us information about the world" (p.7) and "we have very little sense of what a model is in itself and how it is able to function in an autonomous way" (p.8). In Section 2 it is shown that the heterogeneity in modeling practices raises questions on the usefulness of having a single concept that captures so many variegated practices in science. Accordingly, instead of labeling the large number of heterogeneous practices as models it may be more strategic and accurate to avoid this broad categorization which hides rather than reveals the heterogeneity in scientific practices. For example why label metaphors as models if metaphors have a clear and distinct definition? The imperialism of modeling—seeing models everywhere—is, I suggest, impoverishing rather than enriching the study of science. Metaphors are simpler entities and, I shall argue in this paper, can be used to explain a controversy in game theoretic modeling.

As various advocates of metaphors have pointed out

> "When theories run into problems, both the problem and their proposed solutions are consequences of the logic of the metaphors that are at work"[5]

or that

> "Many acrimonious debates in the history of economics would have been clarified tremendously if these [metaphorical] tenets had been kept in view."[6]

and that

> "Controversies in economic discourse can be clarified, and identification slips avoided, if the user of metaphor specifies its type … controversies arising from misunderstanding can be resolved, saving intellectual energy to the defense of the type of resemblance which a particular metaphor is supposed to reveal."[7]

Although not acknowledged by the participants in the controversy, these views suggest that the dispute might indeed be a metaphorical one. The question then is whether this particular controversy is one that should involve, and would be solved by, the identification of strategic uses of metaphors. The question is whether the debate can be framed as one about metaphor choices. The outline of the paper is as follows. Section 2 defines metaphor and discusses its relationships with models and stories. In Section 3 the game theoretic modeling controversy is presented. In section 4 the controversy is cast and assessed as one that involves the choice and strategic use of metaphors. Section 5 concludes.

---

[5] Fernandez-Duque and Johnson (1999:83).
[6] Mirowski (1988:139).
[7] Khalil (2000:7).

1. Why metaphors not models?

.

I argue in this section that metaphors are better understood than models and thus should be used to explain controversies in modeling. Metaphors have had a continuous consistent definition since at least Aristotle. Models are not only a recent tool of reasoning, (the concept of model was barely used before the mid-twentieth century. See Morgan 2012), they are still not very well understood. While metaphors could be described as models, their distinct and simple internal structure would be lost in the large range of heterogeneous practices placed under modeling. If some models have a simple internal structure, the different entities labeled as models is extremely broad.

The received view in economics is that it is a modeling science --modeling is the principle too based mode of reasoning in economics. Mäki (2002:10) remarks "to do economics is to do modeling" and defines, after Robert Solow, model building as a "fact oriented activity that as its objective to isolate key causal dependencies in reality" (p.11). Earlier, Gibbard and Varian (1978) argued as well that economic theorizing consists of investigating economic models. They claim models are used "whenever there is economic reasoning from exactly specified premises" (p.666). Economic models, for Gibbard and Varian, have two elements: stories (that carry the interpretation) and an uninterpreted logical mathematical form (the syntactical part). They argue economics models pose counter-factual questions (of the following sort: what would happen if such and such was the case?) that are useful in generating explanations. Sugden (2002) too emphasizes the explanatory power of models in economics seeing them as describing credible counterfactuals worlds useful to warrant inferences from the model-world to the real-world.

Building on the seminal work of Gibbard and Varian, Morgan (2002) sees models as autonomous agents that contain two modes of reasoning the theoretical and the narrative. Models form a hinge between theory and narrative carving for themselves a relative degree autonomy with respect to both (Morgan and Morrisson 1999). The theoretical parts determines "the relationships between the elements of the model" (Morgan 2002:195), connects the theory to the model, and constrain the narrative. Narrative logic, in turn, connects the model to the economic world. Between the theory and the economic world lies the model as an autonomous complex entity. The question which is left open is what connects the mathematical structure to the narrative? Morgan replies that "where to start the tale, which questions are interesting … is somewhat open –the user has to make sensible choices" (2002:195). This, however, is the function of metaphors—to help connect the mathematical structure to the economic world—to create a mapping between two distinct domains. Morgan (2002) however argues that metaphors and stories are distinct but complement each other in supporting the economist's attempt to explain the world (rather than substitute for each other as for McCloskey (1990) argues). Yet she places metaphors and (mathematical) structures on one side as elements of models arguing that both need stories to produce knowledge about the world (Morgan 2002:183). For her it is stories that connect models to the economic world—this is the function of narrative logic. We must presume then that metaphors and mathematical structures use stories that link them to the economic world.

In the preface of her most recent book on models Morgan (2012:xv) argues that she no longer attempts to offer a definition of models because of the heterogeneity of objects that count as models. She contends models are not easy to characterize and there are no easy answers as to what models are or how modeling works. In fact, "there are lots of different kinds of things that legitimately count as models … and they often look and function very differently" (Morgan 2012:xvi). It seems then that there is no straightforward definition of what models are. Faced with a controversy in game theoretic modeling how then should an economist or a methodologist of economics evaluate it? I will argue in this paper that some controversies could be explained in terms of the strategic use of metaphors. The definition of metaphor as connecting, via a mapping, two distinct domains is all the mental tool kit needed to generate the explanation.

The view adopted here also answers the duality <<the model as a metaphor>> and <<the metaphor as a model>> (Brown 1977). Some see the model as a metaphor (McCloskey, Bronk, Brown), others see the model as an analogy (Klamer and Leonard), and yet others see models as carrier of metaphors (Biccheiri, Morgan). Variegated opinions on what models are is to be expected yet it does not mean there is no agreement how metaphors work. The other side of the duality <<metaphor is a model>> requires explanation as well. This side of the duality appears redundant: An account of how metaphors work can be given without reference to models. As Maasen et. al. (1995:1) explain, in fact, interest lies not in duplicating and expounding "fine grained terminological distinctions between metaphors, images, analogies, models, rhetoric, and systems of thought " but in how metaphors permit "the transfer … of pieces of meaning from one delineable discourse to another".

Economics contains a number of metaphorical expressions including <<Equilibrium>>, <<Elasticity>>, <<Human capital>>, <<Accelerator>>, <<GNP is up>>, <<Prices are inflated>>, <<Liquid assets>>, <<Price mechanism>>, and <<Policy instrument>> (as listed in Klamer and Leonard 1994 and McCloskey 1983). We tend to consider them as thought changing, breaking the habit of thought as it were by employing a deviation from the literal meaning to the figurative meaning: <<Time is money>>, <<Time flies>>, or <<Mind as machine>> (time is not money, time has no wings to fly, and the mind arises from an organic, not mechanic, organ). Metaphors, when not dead, are fluid and open to interpretation being highly sensitive to the context in which they are used. The distinction between the literal and figurative meaning is also not necessarily dichotomous but continuous: expressions which begin as a metaphor harden, freeze, or die of overuse (Bicchieri 1988a; Klamer and Leonard 1994:27). Their death in science, however, unlike poetry, signals a successful metaphor which has spread (Boyd 1979 [1993]).

Another clarification needs to be made here namely on the relationship between analogy and metaphor. For McCloskey (1983) all metaphors belong to the larger class of analogies while Mirowski (1988) uses both terms together and interchangeably. But not everyone agrees. Klamer and Leonard (1994:34), for example, claim analogy and metaphor are close but distinct. In a metaphor they argue there are attributes in common between the *principle subject* (mind, time, market) and the *subsidiary domain* (machine, money, game) whereas analogy is less than a full blown metaphor and requires no imaginative leap since it is based on the Aristotelian principle of proportionality.[8] According to Hoffman et. al. (1990:213), metaphors arise before analogy which are "post hoc relative to the root metaphor". For Klamer and Leonard the analogy is an expanded elaborated metaphor that uses the metaphor to focus on certain relationships in it. One interesting implication is that, for Klamer and Leonard, models are not metaphors but an "explicitly, most often formally articulated analogy" (p.35).

The metaphorical expressions listed above "fill a gap" in the economics lexicon and, while useful, they do not to elucidate how metaphors work in models. Such metaphorical expressions it is said are sanctioned by a higher level metaphorical mapping between two domains (Lakoff 1993:209). I have found six different references that refer to these two domains.[9] In what follows the three that reflect different perspectives on how metaphors work are discussed. The first view posits that metaphorical mappings generate a conceptual mapping of entities, properties, relations, and structure from a *source* to a *target domain* (Fernandez-Duque and Johnson 1999; Lakoff and Johnson 1980; Lakoff 1993). This mapping is illustrated with <<The mind as machine>> metaphor.

[8] In analogy, they further argue, the focus is on the similarities in relationships and (following Aristotle) the principle of proportionality (as a limited and identifiable relationship between subsidiary and principal subjects) suffices.

[9] Klamer and Leonard (1994) mention a few (1) subject / predicate, (2) tenor / vehicle, (3) target /import, and (4) principal/ subsidiary. Fernandez-Duque and Johnson use (5) source / target domain while Bicchieri (1988a) (6) secondary / primary subjects.

The <<Mind as machine>> metaphor[10]

Source domain / subsidiary or secondary subject		*Target domain / principle or primary subject*
Machine	→	Mind
Functions within machine	→	Mental capacities
Products of the machine	→	Ideas
Automated machine functioning	→	Thinking
Normal machine functioning	→	Normal thought
Breakdown of machine	→	Inability to think

Each arrow above takes some entity or structure in the source domain and constructs a counterpart in the target domain. The mappings provide a "fixed pattern of ontological correspondences across domains" that can be activated, not algorithms that mechanically take the source domain inputs and produce target domain outputs (Lakoff 1993:210-14).[11] These mappings also submit to the *invariance principle* which states that "metaphorical mappings preserve the cognitive typology of the source domain, in ways consistent with the inherent structure of the target domain" (ibid.,). The invariance principle can be understood as a constraint on the correspondences that constitute the mapping where cognitive typology is the image-schematic structure of the domains. Similarly, the image-schematic structure of the target domain can in turn limit the possibilities of mappings from the source domain.

A more dynamic account of how metaphors function is given by Black (1962), and earlier Richards (1936), who contend metaphorical meaning arises from the interaction between a *principle* and a *subsidiary subject* or, equivalently, between a *primary* and *secondary subject* (Black 1979 [1993]).[12] While the (unidirectional) mapping gives the principle subject (time or mind) a name that belongs to the subsidiary subject (money or machine), the bi-directional interactionist account entails that the metaphor can modify the prevalent interpretation of both the principles and the subsidiaries as novel meanings, not reducible to or substitutable by a literal expression, arise. The interactionist account captures the

[10] From Fernandez-Duque and Johnson (1999:85).

[11] More specifically, Lakoff (1993:249) explains "contemporary theory of metaphor is at odds with certain traditions in symbolic artificial intelligence and information processing psychology. Those fields assume that thought is a matter of algorithmic symbol manipulation, of the sort done by compute program. This defining assumption is inconsistent with the contemporary theory of metaphor."

[12] Aristotle identified four kinds of metaphors, and though he excluded them from Logic (Klamer and Leonard 1994), his definition appears to be of the interactionist sort since it involves "the transposition of a noun from its proper signification either from genus to the species; or from the species to the genus; or from species to species, or according to the analogous" (Rhetoric, Book III. Chapter XI).

resonance or *expansion* of a metaphor more adequately than the unidirectional view and opens the possibility that the market (as the principle subject that borrowed properties and relations from the subsidiary subject, the game) may in turn also modify our understanding of the game.

Distinctions between various kinds of metaphors have been made. Khalil (2000) identifies four kinds (the nominal, heterologous, homologous, and unificational) while Klamer and Leonard (1994) identify three (the pedagogical, the heuristic, and the constitutive).[13] Pedagogical metaphors are the simplest as they illuminate and clarify a complicated concept by providing mental images to help an audience. They are closest to Khalil's nominal metaphors which only use a superficial similarity between the principle and the subsidiary. They usually help answer the question "what is the intuition?" and "what is the story?" behind a theory or a mathematical model. It is pedagogical metaphors that scientists and economists usually have in mind when thinking of metaphors.

Nominal and pedagogical metaphors, however, are not the most influential in science since they may be omitted without affecting an argument. Heuristic metaphors are part and parcel of theories.[14] They cannot be paraphrased or substituted with a literal expression. Klamer and Leonard argue they are necessary to catalyze our thinking—they are thought propelling. Accordingly, an example of a successful heuristic metaphor is that of <<Human capital>>. The human capital metaphor signals the beginning of an inquiry and will, given the resonance it creates over time, generate additional developments.[15]

Constitutive metaphors frame our thinking, determine what makes sense, and work at the fundamental level of Kuhn's paradigm. They are spectacles necessary for the interpretation of our world and include "those sets of assumptions, usually implicit, about what sort of things make up the world, how they act, how they hang together and, usually by implication, how they may be known … [they] constitute the

[13] Maasen et. al. (1995) identify three kinds of metaphors as well, the illustrative, heuristic, and the constitutive. The definitions of the first two are the same as Klamer and Leonard but their definition of constitutive differs however since they see their function as to replace previous meaning by new ones.

[14] Boyd (1993:486) refers to them as constitutive metaphors but since this term is already employed by Klamer and Leonard to mean something else, it will be avoided.

[15] <<Human capital>> appeared roughly around the same time in Mincer (1958), Schultz (1961), and Becker (1964). Capital in economics classically refers to physical capital such as machinery and plants. Human capital, likewise, referred to education and skills as investments that generate returns for the owner. Human capital is also an input in the production process. The human capital metaphor renders human capital interpretable as one of the inputs (alongside physical capital and technology) in a standard production function. Although an expenditure by individuals, human capital is distinct from the consumption of other goods since it provides a return in the future, like any other investment. In Schultz's case a parallel between physical/capital investments and human skills investment is made explicit. In this way Schultz could use the existing capital terminology to explain large increases in national output. It is here, Klamer and Leonard argue, that the connection between thought in science and metaphor is strongest.

ultimate presuppositions or frames of reference for discourse on the word or on any domain" (Brown 1977:125). Consider the following metaphors <<We've hit a dead-end street>>, <<We can't turn back now>>, <<We may have to go our separate ways>>, <<Look how far we've come>>, <<It's been a long bumpy journey>> etc. They all refer back to the same constitutive metaphor namely <<Love as a journey>> (Lakoff 1993). Because it fitted with the metaphors that constitute neo-classical economics (i.e., physical capital accumulation), constitutive metaphors explain why the <<human capital>> metaphor succeeded as a heuristic metaphor. Another example is the mechanistic world metaphor which generated concepts such as price mechanism, equilibrium, and elasticities among others (Brown 1977; Hodgson 1995). Hard to specify concretely since they operate below conscious awareness and "can be exposed only by digging into or interpreting the relevant texts, both spoken and written" (Klamer and Leonard 1994:41), constitutive metaphors answer the question of where do our heuristic metaphors come from.

While the nominal distinction made by Khalil (2000) overlaps with Klamer and Leonard's (1994) pedagogical metaphor, the heuristic and constitutive metaphors, to the extent that they work for uncovering structures, processes, and powers may overlap as well with Khalil's heterologous, homologous, and unificational metaphors. Khalil, however, is pessimistic about this overlap since he sees Klamer and Leonard not as realists but sophists more interested in persuasion than uncovering real phenomenon. Heterologous metaphors obtain when there is resemblance of analytical function but the context or origins are not the same (for example the wings of a bird and the wings of a bat, although both perform flying, emanate from a different context). As an example of a heterologous metaphor in economics Khalil considers those on spontaneous order arising from climatic and ecological systems (subsidiary) and socio-political order (principle). Homologous metaphors obtain when there is no common analytical function but a similar scheme, context, or common origin (thus though the forelimbs of mice and bats have different functions they have the same origins and are homologous). Examples of homologous metaphors include the evolution and entrenchment of habits and biological evolution; the division of labor within the firm and the differentiation of functions within organisms; and the autocrat of a chimpanzee troop with the modem state. Khalil identifies unificational metaphors as the strongest kind in that it must be the same law operating in the principle and subsidiary subject. Thus the law of gravity is unificational because it is used to explain various physical and astronomical phenomena; the similarity of blood circulation in humans and chimpanzees is also considered a unificational metaphor. As for economics, Khalil claims optimization unifies disparate phenomena by drawing on the similarities between household and firm maximization (utility and profits respectively). Khalil finally explains that the use of any metaphor is appropriate if and only if it is classified in its appropriate category. He provides three levels of identification slips: From single degree (heterologous metaphor is used when the similarity is only nominal) to triple degree (unificational metaphor is used when the similarity is nominal) and offers for each level of identification slip an example taken from economics.

As this section has shown metaphors are more accurately defined than models. All types of metaphors generate a mapping from one domain to another. Equipped with this definition the next section focuses on the backward induction controversy. In section 4 use will be made of the understanding gained in this section to explain this controversy.

2. The backward induction controversy

The economists do not know why they disagree.
McCloskey (1990)

Backward induction (or BI) is a method to compute equilibria in finite, usually perfect information, extensive form games. It involves the analysis of games from back to front proceeding by the elimination of dominated strategies.[16] However, the usefulness and epistemological function of BI led to a controversy to be elaborated on in this section. But rather than use a model to explain this controversy in the heart of modeling practices in economics, this paper uses metaphors. The question to be asked is in what terms this controversy can be explained? Rather than use a model to explain a controversy in modeling, and risk reflexivity, use is made of a different mode of reasoning, the metaphor.

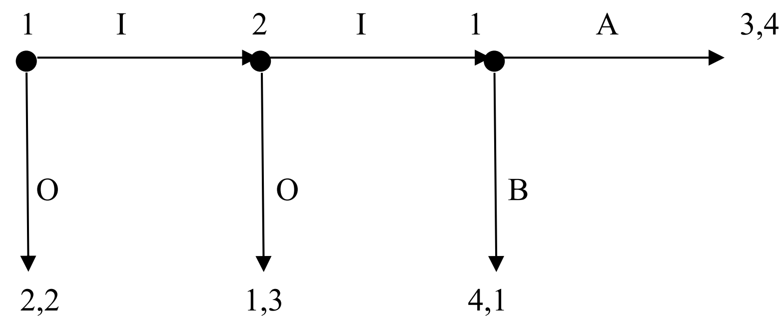
BI led to the Nash refinement literature and to so-called subgame perfect equilibrium (Selten 1975). The term refinement is used since such equilibria involve additional criteria which take the form of eliminating non-credible threats. Subgames require an initial node and are self contained games in a larger game (including the game itself). A node $\boldsymbol{x}$ initiates a subgame if neither $\boldsymbol{x}$ nor any of its successors are in an information set that contains nodes that are not successors of $\boldsymbol{x}$; the subgame is the tree structure defined by such a node $\boldsymbol{x}$ and its successors (Watson 2013). A subgame perfect Nash equilibrium obtains when there is a Nash equilibrium in every subgame of the larger game and when a subgame is reached,

[16] The first explicit reference to BI is due to Kuhn (1953). Luce and Raiffa (1957:68), though not the first, state that "at a terminal choice point—we are assuming that all games have a stopping rule, and this enables us to work backward—the player whose move it is will naturally adopt the choice which suits him best. Thus, since the last choice is determinate, we may as well delete it and place the appropriate payoff directly at the terminal move position, if this is done for each terminal move, the penultimate moves now play role terminal moves, and so the process may be carried backward to the starting point."

the players will play according to this equilibrium strategy. It follows that not all Nash equilibria are subgame perfect but a subgame perfect equilibrium has to be a Nash equilibrium.[17] This classical account of playing games (to be further described below) is, via BI, considered "the only possible pattern of play by rational players" (Biccheiri 1988b:383).

BI combined with rationality and common knowledge of rationality (CKR) entails that in games such as the centipede players choose at every decision node "down" in Figure 1 until the first node is reached.

Figure 1. Rosenthal's (1981) extensive form centipede type game (player 1, player 2).

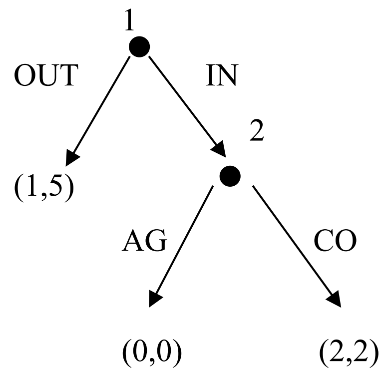


To understand why, note that in the Figure 1 player 1 will opt for down (B) in the last node of the game because the payoff is highest (4>3). Knowing this player 2 plays down as well since this maximizes payoff (3>1). Back to the first node of the game, player 1 knowing player 2 will play down if she had the chance will play O (2>1) ending the game before it starts. In repeated games, players play various rounds (of the same game). BI again implies non-dominated choices are made: Starting from the last game each player opts for the choice that maximizes own pay-off irrespective of what the other plays until the first game is reached. In the finitely repeated prisoner's dilemma this means that players will defect on each and every round of the game. The repeated prisoner dilemma, although not a perfect information game, thus submits to a similar reasoning and both players are expected to defect. A third game where BI is applied is the chain-store game in Figure 2 (Selten 1978). Here an incumbent monopolist (player 2) holds a monopoly in *m* towns. The monopolist does not declare a price war when one of the *m* competitors, one in each town, decides to compete. In the chain store game the monopolist's decision to play AG (for aggressive or price war) is seen as a non-credible threat since it is not in its interest to retaliate and it will play CO (for cooperate) (Gibbons 1997). It should be noted here that this game has two Nash equilibria {IN, CO} and {OUT, AG} but only not declaring a price war {IN, CO} is subgame perfect.

[17] This is because all non credible threats Nash equilibria have been eliminated from the smaller set of subgame perfect Nash Equilibrium. Simulations have also shown that it is possible to derive Nash equilibria that are not subgame perfect (Binmore and Samuelson 1996; Gale, Binmore, and Samuelson 1995).

Figure 2. Selten's (1978) Chain store game in extensive & normal form

1 \ 2	CO	AG
IN	2,2**	0,0
OUT	1,5	1,5*



*Nash Equilibrium

**Sub Game Perfect Nash Equilibrium

Selten (1975:35) early on identifies a difficulty with BI noting "there cannot be mistakes if the players are absolutely rational". Selten (1978) also observes that there is a paradox because while it is more advantageous for the monopolist to cooperate in the short run (the BI and game theoretic correct choice), it is better for him to play aggressive in the long run (the deterrence, more convincing theory). Accordingly, the repeated chain store game is paradoxical because only BI is theoretically correct yet playing aggressive—and starting a price war-- is much more convincing. Nevertheless, Selten argues that a satisfactory interpretation of equilibrium in extensive games seems to require the possibility of mistakes is not completely excluded. Selten introduces irrational play assuming players are subject to rationality imperfections so that at every information set u there is a small probability ε_u for the breakdown of rationality. He sets the stage for the controversy noting that there cannot be any unreached information set and that this is consistent with the definition of strategy profiles which inform what the player will do at every information set of the game, thus specifying behaviors even over unreached subgames.

Experimental results on the use of BI by actual players in games are ongoing but far from conclusive. In over 1000 experiments conducted since the late 1950s cooperative choices were made in about 30% of the time in the repeated prisoner dilemma (even as it emerged that experience raises the chances of defection or playing the BI equilibrium; see Andreoni and Miller 1993; Colman 1998:356). Binmore, Shaked, and Sutton (1985) find that in a bargaining game individuals learn to play BI in round two when the roles are reversed. Balkenborg (1998) runs experiments over what he calls the basic (or stage) game and notes that 80% of the results support the outcome predicted by BI (with 13 sessions and 12 subjects each playing the game 50 times randomly varying with anonymous opponents). Johnson et. al. (2002) test the extent to which deviation from the BI path is explained by "limited cognition" or "equilibrium social preferences". They find that both contribute to explaining deviations from BI and suggest that individuals are not equipped to use BI without prior training. After parceling the ultimatum game[18] into rationality, subgame consistency, and truncation consistency Binmore et. al. (2002) find evidence against the use of BI. They thus back the long held result of Güth et.al. (1982). There are also experimental results against the use of BI in the ultimatum game (Henrich et. al., 2005; Guala 2008; Roth et. al. 1991), the *p*-beauty contest (Camerer 2003a), and the centipede game (McKelvey and Palfrey 1992).

A common critique against BI is that it is paradoxical (Basu 1990; Biccheiri 1989; Luce and Raiffa 1957; Pettit and Sugden 1989). The culprits are the conjunction of CKR and rationality. The argument is that if players are rational they may have to consider cooperation. Yet, and here lies the paradox, BI renders cooperation non-utility maximizing as it assumes (in the centipede game for example) player 2 plays down if the second node, which is not meant to be reached, is reached. If node 2 is not meant to be reached according to the theory how can it predict that player 2 plays down? Pettit and Sugden's (1989) argument against the BI equilibrium is innocuously simple and follows this train of thought, namely, that common belief in rationality breaks down--and the BI equilibrium does not obtain--when one of the players acts cooperatively in the repeated prisoner's dilemma. The paradox was demonstrated more rigorously in Reny's (1993) proof that rationality (as utility maximization) and CKR are inconsistent in two person perfect information finite games. If either CKR or rationality is dropped, Reny further argues, BI is no longer the only type of rational play. Reny's own interpretation of his proof questions the plausibility of subgame perfection.

Until now the BI equilibrium has been criticized on at least two connected grounds: It is not reasonable for players not to cooperate when this will benefit all with enough rounds to play the game; and players must take hypothetical decisions in nodes they will never reach (CKR and rationality are inconsistent). The controversy, however, really begins after Aumann's (1995) formal mathematical proof that CKR and rationality in a perfect information game suffice to justify the BI equilibrium. More specifically, Aumann argues that when CKR and rationality obtain no vertices off the BI path are reached (p.18). Aumann sustains that the proof supports the intuition that common knowledge implies BI while relying on the usual meaning of concepts such as knowledge and rationality. Irony has its ways since it was mostly as a

[18] In the ultimatum game for example player 1 makes an offer equal to a portion of a total sum to be shared with a receiver who then decides to accept or reject the offer. If the offer is rejected—as often happens when the offer is much under 50% of the total-- no players gets any payoff. If the offer is accepted both players get the agreed sums. The subgame perfect equilibrium is for the receiver to accept any positive offer that is made since something is better than nothing. Rejection rates of positive non-trivial offers however are quite common across many cultures.

response to Reny et. al. that Aumann (1995) offered his proof. At this stage perplexion is allowed since we are left with a proof against a proof (Reny vs Aumann) on the one hand or common sense and intuitions against the proof on the other (Aumann vs Pettit, Sugden, Biccheiri, Basu, Selten et. al.). Even after Aumann (1995) published his proof, both supportive (Rabinowicz 1998) and contesting counter-proofs (Binmore 1997; Ben –Porath 1997) were published.[19]

Binmore and Samuelson (1996) were quick to counter Aumann (1995), initially disputing the use of mathematical logic to establish the proof. Binmore (1997:24) subsequently adds that the proof reduces to "inventing fancy formalisms … only to confuse matters". Aumann's proof it was also claimed brushes aside the question of how the players acquired rational and unambiguously defined beliefs in the first place. For Binmore and Samuelson, equilibria should not be established via the static definition-axiom-theorem-proof format (that closes the mind) but via algorithms of players' reasoning and "constructive" simulations of the equilibrating process.

According to Binmore and Samuelson, the problems of classic game theoretic rationality are compounded by CKR. Unless counterfactuals such as rational players acting irrationally are accounted for in the strategy profiles of players, rationality makes no sense. According to the critique, the BI path can only be justified with counterfactuals (or choices that could have been made but were not).[20] The possibility of hypothetical decisions of the sort "what should player 2 do if player 1 does not play down but cross instead?" is now seen as requiring special attention. How would a player explain that the other player is not playing the BI path? What kind of mistake or irrationality brought us to node 50 out of 100 instead of ending the game at the first node? These questions were raised by Binmore and Samuelson (1996) and Binmore (1996; 1997) as a critique of Aumann (1995). While the critique acknowledges that rational play leads the first movers to play down (eg., Biccheiri 1988b; Binmore 1987:196;1997), the critique also states that rationality is not adequately modeled if players do not account for what they would have done if the other player does not follow the BI path.

---

[19]Without CKR, Rabinowicz defends BI for a class of BI-terminating games where rationality is *at* choice of moves not strategies. Here BI-terminating games are games such as the centipede games where down ends the game excluding the finitely repeated prisoner dilemma. Binmore uses a finite version of the centipede game to show that even with CKR the equilibrium of the game is a Nash mixed strategy equilibrium, not the BI equilibrium. Ben-Porath assumes CKR only at the first node, exploiting the distinction between certainty which allows surprises (playing cross with probability 0) and knowledge which does not, as well as the possibility of changing beliefs, Ben-Porath claims that the BI equilibrium is no longer the only justifiable equilibrium.

[20] Note that such arguments require observable behaviors by players or else there would not be a paradox and indeed there is no such paradox in simultaneous move games (Reny 1993).

Aumann's proof was meant to answer these critiques which turned only more virulent insisting his conception of rationality is mistaken as long as it does not specify what the players would play if they deviate from the path of BI. Aumann's (1996a,b) response is that his proof (i) does not necessarily imply rational players will not deviate from the BI path; (ii) he insists that rational players may deviate at any point including the first move; and (iii) that the inductive choice used in the proof could be irrational. He acknowledges that his conception of rationality and strategy does account for player *i*'s knowledge of what the other player would do had *i* played across instead of down (in the centipede game). Aumann suggests his critiques are confusing the assumption of rationality with CKR and shows that his theorem still holds if he adopts Binmore's stronger definition of rationality. It is relevant to add here that Aumann's interpretation has recently received the support of an empirical game theorists namely, Gintis (2009). In summary, Aumann (1996a,b) argues that as long as rationality is common knowledge, a player picking across is absurd logically as down is the only possible outcome at the first information set of the game.

By the looks of it, the debate has stalled with both sides entrenched and unable to dig deeper for answers. The core of the controversy is what (if anything) can justify BI? For skeptics, when rationality and CKR do not contradict each other, something else beyond rationality and CKR, is needed, including maybe stories from outside the game. As tension emerges between game theoretic reasoning on the one hand and intuition or common sense on the other hand, is it surprising that in the most formal modeling branch of economics—game theory—so much weight is given to intuition? If metaphors are as important as their proponents argue they are this should not come as a surprise. The presence of intuitions and common sense may still be explained as an inchoate metaphorical choice that needs to be formally acknowledged. The literature on metaphors can step in to formally account for these choices as done next.

4. The strategic use of metaphor

How can use be made of metaphors to provide an explanation for the controversy on BI, rationality, and CKR in game theory? Recall Aumann assumes rationality and CKR sufficient to justify the use of BI in computing equilibria. This sufficiency relies on a mathematical proof that shows irrational choices could not have been made (Aumann 1996a,b). Binmore and Samuelson (1996) on the other hand disagree and posit that his mathematical proof is irrelevant: Something else is needed to justify BI. The nature of the controversy between Binmore and Samuelson, on the one side, and Aumann, on the other, remains puzzling and in need of explanation. Indeed, why is there a controversy if (i) the mathematical proof of BI with CKR is rigorous?[21] And (ii) if empirical tests unfavorable to BI could but are not brandished to undermine its validity? Binmore and Samuelson (1996) argue that the controversy is not within mathematics: They (albeit not others) agree the mathematical proof that CKR leads to the BI equilibrium is rigorous yet they suggest that the proof has little value. The critiques of Aumann are convinced something is not right, that the BI equilibrium cannot be justified the way Aumann does (In fact his way is perceived as so fundamentally wrong by Binmore that it could lead game theory stop being taken seriously).

This section uses the definitions of metaphors provided earlier to explain the confusion around the status of BI. The first, if obvious, possibility is that Aumann uses a constitutive metaphor different from the one adopted by his critiques. Keeping in mind that constitutive metaphors, as noted by Klamer and Leonard, are hard to specify concretely, we can posit that Aumann's constitutive metaphor casts the foundations of game theory in mathematical logic which is the source domain of its theoretical results in terms of solution concepts and equilibria. Within such a constitutive metaphor there is little sensitivity to the context or type of game. Indeed most of Aumann's results can be applied to a range of games, not one. Here there is less concern for the structure and context of the game than his critiques would like. In mainstream economics there could be other competing and conflicting constitutive metaphors behind the critique's claim that game theory may become irrelevant if it adopts too abstract or idealized foundations à la Aumann. The critique operates under at least one, possibly various, constitutive metaphors that do not use the same source domain as the foundation of game theory.

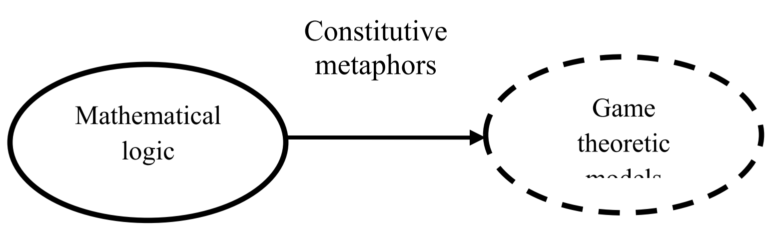
If this controversy is indeed about the adoption of different constitutive metaphors it should be acknowledged as such. The confrontation accordingly is not over who has the best mathematical proof or which theory is more supported empirically but which constitutive metaphor is more adequate for the foundations of game theory. Focusing on constitutive metaphors also forces questions on explaining why different constitutive metaphors are employed by different individuals. While helpful in some ways, framing the controversy simply as one of constitutive metaphors fails to provide a comprehensive explanation for the controversy. For a start it would force us to place Binmore, Samuelson, or Biccheiri under the spell of one constitutive metaphor, a highly doubtful claim. Indeed, there could be various constitutive metaphors uniting the opponents of Aumann on what the foundations of game theory ought to

[21] Binmore (1987:196), for example, claims "It is not disputed that the results of the play of this [centipede] game by rational players will be that I plays "down" at the first node." Binmore (1997) also provides his own proof in (1997).

be based on. Second, the conflict of constitutive metaphors does not rule out the possibility that some of Aumann's critiques are with him in the same constitutive metaphor of mathematical logic (eg., Reny 1993). And third among those who defend the equilibrium of BI not all use exacting mathematical logic "that closes the mind" but stories and common sense arguments (eg., Broome and Rabinowicz 1999; Sobel 1993). Unless there is a yet unidentified constitutive metaphor which could tightly explain by regrouping those for and against the justification of BI via common knowledge of rationality, constitutive metaphors for the time being are not sufficient to explain the controversy away.

This does not mean constitutive metaphors play no role in the controversy. A finer possibility is to frame the controversy as one of three metaphorically interconnected but relatively autonomous domains namely mathematical logic, the game, and the economic world (See Figures 3-5). A similar division is used by Grüne-Yanoff and Schweinzer (2008) to describe the architecture of game theory. Their view is as ours but lacks the metaphorical theoretical justification that explains the triptychal architecture. Game theory, connects or bridges the gap between the domains of mathematics and the social world. It is a semi-autonomous model between mathematical logic and the complex economy (Morrison and Morgan 1999). Each source domains maps its properties into game theory. Accordingly, mathematical logic and the economy map theirs properties into game theory which then generates game theoretic models. The metaphors on the right hand side that connect the model to the economy are much wider than the unificational metaphor that connects mathematical logic to the model because the model and more so the world are highly complex entities.

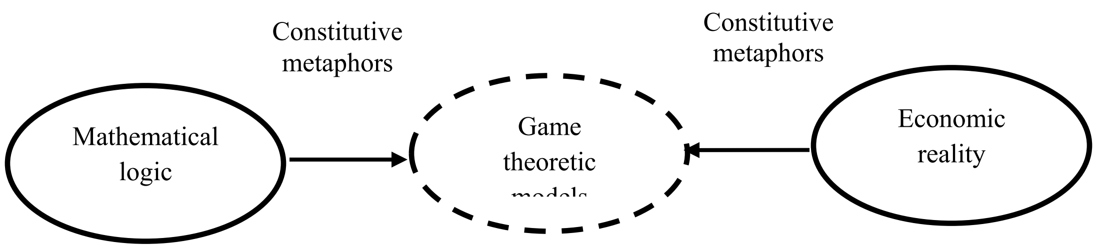
Figure 3. Aumann's constitutive metaphor



Only unification metaphors allowed

Aumann's practice of game theory is limited to the constitutive metaphor in Figure 3. His metaphor allows only unificational mappings that connect mathematical logic to game theory. This is where Aumann is mostly active (in the controversy at least): he maps mathematical logical properties into game theory, constructing different models that say nothing specific about the right hand side, the economic world. A wider constitutive metaphor, one that allows various kinds of heuristic, unificational, heterologous, and homologous metaphors operate between the game and the economic world on the right hand side. Binmore, Samuelson, and others partake in the controversy using metaphors from both ends (Figure 4). Accordingly they do not privilege metaphors from the left hand side only. While they use some left hand side mapping as well, they also focus on metaphors that hook the theoretical models of game theory to the world.

Figure 4. Binmore and Samuelson's metaphorical mappings



Heuristic, unificational, heterologous, and homologous metaphors allowed

The changing metaphors across the triptych can lead to identificational slips of the form identified by Khalil, or to the breaches of invariance identified by Lakoff. Such slips and breaches are explained by the prevalent understanding of what games are and what they are attempting to do. They may be strategically used to undermine the work of an opponent. Binmore (1997) for example freely uses proofs—and even has his own version of the BI equilibrium proof—yet he at the same time criticizes Aumann's use of mathematical logic to prove the BI equilibrium. Binmore and Samuelson's slip is to project heterologous or homologous metaphors that operate between the game and the economic world (via simulations and computational economics) onto the unificational relationship that connects the game to mathematical logic. In the language of the Lakoffian invariance principle, their metaphorical mapping violates the image-schematic structure mathematical logic maps into the game, its target. The acceptance of this identificational slip or breach of invariance is of course contingent on accepting our description of the controversy as a metaphorical triptych between mathematics, the model, and the economic world.

Within each of the constitutive metaphors, the participants in the controversy employ distinct heuristic, heterologous, homologous, or unificational metaphors. Since Aumann (1995) is working on the left side of the triptych his proof is consistent with the constitutive metaphors that unifies game theory and mathematical logic. Such unificational metaphors have the principle as the game and the mathematical proof as subsidiary. The connection is unificational because the same principles of mathematical logic obtain in the principle and the subsidiary subjects. The properties of the mathematical proofs are mapped as possibilities of solution concepts in the game (they constrain the theoretical form of the game in the same way the structure of the game constrains the metaphors that connect it to the economic world). One such proof uses CKR and rationality to explain the game, or the equilibrium of the game, via BI. When Aumann notes that he wants to keep the proof as transparent and as simple as possible he echoes Biccheiri that familiarity and manageability are two parameters that guide the selection of metaphor. Aumann is comfortable with the principles of mathematical logic (his subsidiary subject is both familiar and manageable) to prove that the equilibrium entails the first player plays down. Furthermore, Aumann (1995:6) suggests that this proof is expansionary since there is a modern refinement literature supporting it and his own work extends a number of recent papers on similar fundamental notions of noncooperative game theory. Aumann's interpretation of the proof has received support from various quarters including applied game theorists such as Gintis (2009).

Binmore and Samuelson are more aware that they are dealing with metaphorical choices and acknowledge that (i) context matters especially when nodes not meant to be reached by rational players are reached (Binmore 1987:196); and that (ii) a preliminary informal classification of different equilibrating processes through the choice of interesting environments in which games are played should be made (Binmore 1987:183). As noted in Section 2, context is critical for metaphors to be rigorously interpreted. Aumann, however, suggests his proof is context free and works as an ideal gas whose implications are not affected by what happens in practice. Statements by Binmore (1997:28) that "it is at the interpretive level that the importance of common knowledge assumptions needs to be acknowledged") are significant for choices of metaphors. Binmore and Samuelson are explicit that Aumann's subsidiary domain (mathematical logic) is inadequate to justify the computation of equilibria. They claim that the current sequence of "axiom-definition-theorem-proof" does not just close the mind to irrelevancies (a good thing) but also closes it to issues it is perilous to neglect. They suggest instead that the proper way to identify potential equilibria is the use of simulations and stories to interpret counterfactuals. Such modes of reasoning are sanctioned by the broad constitutive metaphor that connects the model to the economy.

Binmore (1987) presents his approach to game theory as fundamentally different from Aumann's. He, first, posits an algorithmic, "machine programmable", definition of rationality.[22] Second, Binmore and Samuelson (1996:114) proceed to search for stories that explain deviation from the rational BI path.[23] Binmore and Samuelson have plenty of stories to overcome Aumann's problem which is seen as the traditional approach to game theory. Their approach identifies mechanisms that explain deviation from rational play (trembling hand, irrational mistakes, defective reasoning). What is the relationship between stories and metaphors? The stories use metaphors to form subsidiary subjects whose properties are mapped onto the principle subject , the solution concept of the game that needs explanation. As Morgan (2007:169 emphasis added) states, for the prisoner's dilemma, "The narratives translate the prisoners' situation into the economic situation—*they link particulars to particulars*—and "explain" how it is, for example, that two large firms can end up doing damage to each other just as the prisoners end up with the double defect outcome". The constitutive metaphor sanctions a particular type of story which uses a particular type of heuristic, unificational, heterologous, or homologous metaphor to connect distinct domains. The metaphor and the story thus complement each other.

Two of the stories identified by Binmore and Samuelson provide counterfactuals without which rationality and deviation cannot be accounted for. Selten (1978), on the one hand, posits deviations from the BI path as a trembling hand which makes mistakes. Kreps et. al., (1982), on the other hand, focus on modeling incomplete information into the finitely repeated prisoner dilemma. Kreps et. al., posit that when a player is not sure what the other will play and there is a small positive probability ($\delta>0$) she may cooperate, tit-for-tat cooperation will be the equilibrium outcome for long periods in finite games, including the prisoner's dilemma. A third possibility, not discussed by Binmore and Samuelson, but suggested by Coleman (1998), is to substitute rationality (and therefore modify CKR) with non-monotonic reasoning that reflects common sense, everyday, reasoning. This modification, Coleman argues, solves the BI paradox by offering theoretical options for players that face the unexpected cross play in the centipede game, cooperation in the repeated prisoner's dilemma, or the declaration of a price

[22] For Binmore (1987:181) a rational decision process refers to the "entire reasoning activity that intervenes between the receipt of a decision stimulus and the ultimate decision, including the manner in which the decision-maker forms the beliefs on which the decision is made. In particular to be rational will not be taken to exclude the possible use of the scientific method."

[23] Their alternative interpretation of the centipede game involves a husband who, after missing his mortgage, explains to his (furious) wife that he would not have lost the repayment had he been dealt the Ace of Diamonds rather than the Queen of Spades in last night's poker game. They point out such counterfactual stories are not obtained from abstract mathematical contemplation but as stories from the world. In the centipede game Binmore and Samuelson discuss two possible competing interpretations of counter-factuals. Both stories provide an explanation of irrational play, or play that plays cross not down or out.

war to prevent further entry in the chain-store game. All three stories provide entry points for distinct metaphors.

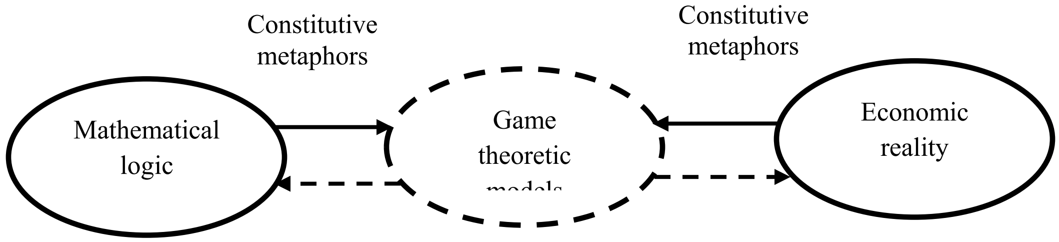
Who has the better story-metaphor combination? Binmore and Samuelson consider this is a wrongheaded question which has no absolute answer since it is always necessary "to look at the context in which the game is played for inspiration on this score. But this context is exactly what is abstracted away when one adopts the conventional mathematical formalism" (p.115). Thus the trembling hand story is not applicable to chess—it provides a poor mapping from one domain to another (a chess player is unlikely to consistently make the same mistake due to a trembling hand). Binmore (1987) accordingly states that irrational play in games such as chess should be modeled--not as a trembling hand mistakes--but as defective reasoning. Which metaphor provides the most adequate mapping depends on the type of game and the context of application.

Understanding the dynamic interdependence-autonomy between the three domains of game theory can shed some light on its evolution. Whether the constitutive metaphors are fixed or not, whether one can take over the other, is an open empirical question. The triptych also depicts how game theory can performatively change both mathematical logic and the social world (dashed arrows from game theory to mathematical logic and the world; Figure 5). An example of this possibility is provided by Morgan's (2007:159): analysis of a Second World War text by game theoretician Rapoport (see also Rubinstein 2006)

> While Rapoport suggests game theory was taken because of the "civilization" of war, it seems equally part of the process that war became acceptable because it was reinterpreted in game theory terms … the cold war came to be seen as a set of game situations … it comes to the point at which we understand and interpret that [nuclear arms] race as a prisoner's dilemma game.

Properties of the model are mapped onto the world via game theory--the model changes the way the world is perceived ("the cold war came to be seen as a set of game situations"). But there is another metaphorical loop that maps properties of the models into mathematics pushing for novel interpretations in mathematical logic and leading to developments in mathematics (mathematics as an applied science that can experience empirical discoveries and novel interpretations). Accordingly, the generation of new game theoretic models can lead to the development of novel mathematical objects, theories, and techniques.

Figure 5. Performativity and game theory



The last point to be made here refers back to the BI micro industry mentioned in the introduction. I distinguish between pre- (Basu 1977; 1990; Biccheiri 1988b; 1989; Binmore 1987; Pettit and Sugden, 1989; Reny 1993; Selten 1978; Sobel 1993; Sugden 1992) and post- Aumann (1995) publications (Ben-Porath 1997; Broome and Rabinowicz 1999; Aumann 1996a,b; 1998; Binmore 1996; 1997; Binmore and Samuelson 1996; Rabinowicz 1998; Stalnaker 1996). Morgan (2007:176) proves helpful here again with her contention that game theory has grown "from the narratives, which … go through a process of matching the economic situation with the game situation and then exploring how and why it does not fit. When it does not fit, a new version of the game is developed with slight changes in the rules, payoffs, or information arrangements". A snapshot review of only a few specimens from the post- Aumann (1995) publication support Morgan's diagnostic. Stalnaker (1996), for example, posits common *beliefs* of rationality instead of CKR to defend the BI equilibrium (while Sugden (1992), pre- 1995, does the opposite, namely, he uses so-called entrenched common beliefs to overcome the paradox of CKR and BI); Rabinowicz (1998) defends BI for a class of BI-terminating games where rationality is at choice of moves not strategies; Ben-Porath (1997) assumes CKR only at the first node, exploiting the distinction between certainty which allows surprises (playing cross with probability 0) and knowledge which does not; Aumann (1998) distinguishes between ex-ante and ex-post knowledge operators of rationality and argues that the proof of the BI equilibrium in the centipede game via (a less intuitive) ex-ante definition of rationality subsumes (a more intuitive) ex-post definition of rationality. And so on and so forth. For a few exceptions (cf., Camerer 2003b and Colman 1998) this micro industry on the BI controversy does not empirically confront BI, CKR, or rationality. Instead, it creates interminable new taxonomies based on changing assumptions, introducing new definitions, logical proofs, lemmas, and theorems. Though the author's view on the value of this micro industry is secondary, it is not clear where lies the epistemological contribution to the social sciences.

5. Conclusion

This paper considered the efficacy of the strategic use of metaphors to explain a controversy in game theory. It suggests that, appropriately employed, metaphors shed light on the possible source of the BI controversy and that the disagreement may be due to operating under a different metaphorical spell. The economists involved in this controversy published past each other over years possibly because the source of their misunderstanding, and that which would have aligned the discussion planes, is an acknowledgement of the strategic use of metaphors. While economists trade is model building model based reasoning has not led to the breakthrough that would have reconciled their differences. The metaphor was introduced as a simpler and possibly looser mode of reasoning to explain this controversy. The metaphorical explanation -- if adopted--would move the controversy to a common plane focusing the discussion on the context and source of their disagreement. Even if the disagreement does not vanish and protagonists stick to their guns as it were, at least they would have understood its source.

Binmore's and Samuleson's critique of Aumann led us to consider various possible uses of metaphors. At the highest level, the level of the paradigm as it were, the confrontation is over which constitutive metaphors is more adequate not which proof defeats the other proof or the extent to which proofs are weakened by empirical evidence or common sense play. Other metaphors within the theme of the constitutive metaphors could be operating. Their success has been linked to their uptake and expansion in the discipline. Uptake and expansion remain, however, insufficient success criteria for an advocate of realism in the sciences. Contributions that explicate the necessary function of metaphors for realism such as Boyd's or Khalil's can step in here. Realism, nevertheless, appears to be more relevant over the metaphorical mapping that connects games to the world, not games to mathematical logic.

Finally, our account of the controversy entails the left side of the triptych (that maps mathematical logic into the game) is not conceptually insulated from the economic world, rather, any change there is to be mediated by the game, an evolving semi-autonomous structure not reducible to mathematical logic nor expandable to the economy. Potential change is a complex and negotiated endeavor over constitutive metaphors and a variety of other none game theoretic considerations. Here the constitutive metaphor acts as a filter. Unless the changes are consistent with it they may not even be considered as serious candidates and will be rejected for committing identification slips and breaches of the invariance principle. Binmore and Samuelson's attack on the "axiom-definition-theorem-proof" sequence could be an example of an identificational slip that challenges the prevalent constitutive metaphor that posits a unificational relationship between game theory and mathematical logic.

6. References

Amadae, S. (2008) *Rationalizing Capitalist Democracy*. Chicago University Press.

Andreoni, J., and Miller, J. (1993) Rational cooperation in the finitely repeated prisoner dilemma: Experimental evidence. *Economic Journal*, 103, pp. 570-585.

Aumann, R. (1995) Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8, pp. 6-19.

Aumann, R. (1996a) Rejoinder. In *The Rational Foundations of Economic Behaviour*, Proceedings of the IEA Conference held in Turin, Italy, Enrico Colombatto, Mark Perlman, Christian Schmidt (eds). Palgrave Macmillan.

Aumann, R. (1996b) Reply to Binmore. *Games and Economic Behavior*, 17, pp. 138-146.

Aumann, R. (1998) A note on the centipede game. *Games and Economic Behavior*, 23, pp. 97-105.

Balkenborg, D. (1998) An experiment on forward versus BI . Working Paper.

Basu, K. (1977) Information and strategy in iterated prisoner's dilemma. *Theory and Decision*, 8, pp. 293-198.

Basu, K. (1990) On the non-existence of a rationality definition for extensive games. *International Journal of Game Theory*, 19, pp. 33-44.

Becker, G. (1964) *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*. Chicago: University of Chicago Press.

Ben-Porath, E. (1997) Rationality, Nash equilibrium, and backwards induction in perfect information games. *The Review of Economic Studies*, 64, pp. 23-46.

Bicchieri, C. (1988a) Should a scientist abstain from metaphor? In *The Consequences of Economic Rhetoric*, Arjo Klamer, Donald McCloskey, and Robert Solow (eds). Cambridge University Press.

Bicchieri, C. (1988b) Common knowledge and BI : A solution to the paradox. Behavioral Ethics Lab, University of Pennsylvania.

Bicchieri, C. (1989) Self-refuting theories of strategic interaction: A paradox of common knowledge. *Erkenntnis*, 30, pp. 69-85.

Binmore, K. (1987) Modeling rational players I. *Economics and Philosophy* 3, pp. 179-214.

Binmore, K. (1996) A note on BI. *Games and Economic Behavior*, 17, pp. 135-137.

Binmore, K. (1997) Rationality and BI . *Journal of Economic Methodology*, 4, pp.23-41.

Binmore, K., and Samuelson, L. (1996) Rationalizing BI . In *The Rational Foundations of Economic Behaviour*, Proceedings of the IEA Conference held in Turin, Italy, Enrico Colombatto, Mark Perlman, Christian Schmidt (eds). Palgrave Macmillan.

Binmore, K., Shaked, A., and Wolinsky, A. (1986) Testing non-cooperative bargaining theory: A preliminary study. *American Economic Review*, 75, 1178-80.

Binmore, K., McCarthy, J., Ponti, G., Samuelson, L., and Shaked, A. (2002) A BI experiment. *Journal of Economic Theory*, 104, pp. 48-88.

Black, M. (1962) *Models and Metaphors*. Ithaca, NY: Cornell University Press.

Black, M. (1979 [1993]) More about metaphor. In *Metaphor and Thought*, Andrew Ortony (ed). Cambridge University Press.

Boumans, M. (1999) Built in justification. In *Models as Mediators*, Mary Morgan and Margaret Morrison (eds). Cambridge University Press.

Boyd, R. (1993) Metaphor and theory change. In *Metaphor and Thought*, Andrew Ortony (ed). Cambridge University Press.

Bronk, R. (2008) *The Romantic Economist. Imagination in Economics*, Cambridge University Press.

Broome, J., and Rabinowicz, W. (1999) Backward induction in the centipede game. *Analysis*, 59, pp. 237-242.

Brown, R. (1977) *A Poetic of Sociology. Toward a Logic of Discovery for the Human Sciences*. University of Chicago Press.

Camerer, C. (2003a) Behavioral studies of strategic thinking in games. Trends in *Cognitive Sciences*, 7, pp. 225-231.

Camerer, C. (2003b). *Behavioral Game Theory*. Russell Sage Foundation and Princeton University Press.

Clarke, J., Holt, R., and Blundel, R. (2014) Re-imagining the growth process: (co) evolving metaphorical representations of entrepreneurial growth. *Entrepreneurship and Regional Development, An International Journal*, Online First, pp. 1-23.

Coleman, A. (1998) Rationality assumptions of game theory and the backward induction paradox. In *Rational Models of Cognition*, M. Oaksford and N. Chater (eds). Oxford University Press.

diSessa, A. (1993) Towards an epistemology of physics. *Cognition and Instruction*, 10, pp. 105-225.

Emmeche, C., and Hoffmeyer, J. (1991) From language to nature—the semiotic metaphor in biology. *Semiotica*, 84, pp. 1-42.

Fernandez-Duque, D., and Johnson, M. (1999) Attention metaphors: How metaphors guide the cognitive psychology attention. *Cognitive Science*, 23, pp. 83-116.

Fisher, I. (1911) *The Purchasing Power of Money*. New York: Macmillian.

Gale, J., Binmore, K., and Samuleson, L. (1995) Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8, pp. 56-90.

Gibbard, A., and Varian, V. (1978) Economic models. *The Journal of Philosophy*, 75, pp. 664-677.

Gibbons, R. (1997) An introduction to applicable game theory. *Journal of Economic Perspectives*, 11, pp. 127-149.

Gintis, H. (2009) *The Bounds of Reason*. Princeton University Press.

Grüne-Yanoff, T., and Scheinzer, P. (2008) The role of stories in applying game theory. *Journal of Economic Methodology*, 15, pp. 131-146.

Guala, F. (2008) Paradigmatic experiments: The ultimatum game from testing to measurement. *Philosophy of Science*, 75, pp. 658-669.

Güth, W., Schmittberger, R., and Schwarze, B. (1982) An experimental analysis of the ultimatum game. *Journal of Economic Behavior and Organization*, 3, pp. 367-368.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N., Hill, K., Gil-White, F., Gurven, M., Marlowe, F., Patton, J., Tracer, D. (2005) "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28, pp. 795-855.

Hodgson, G. (1995) Biological and Physical Metaphors in Economics. In *Biology as Society, Society as Biology: Metaphors*, S. Maasen, E. Mendelsohn, and P. Weingart, P. Kluwer Academic Publishers.

Hodgson, G. (2005) Decomposition and growth: Biological metaphors in economics from the 1880s to the 1980s. In *The Evolutionary Foundations of Economics*, K. Dopfer (ed.). Cambridge: Cambridge University Press.

Hoffman, R., Cochran, E., and Nead, J. (1990) Cognitive metaphors. In *Metaphors in the History of Psychology*, D. Leary (ed). Cambridge University Press.

Johnson, E., Camerer, C., Sen, S., and Rymon, T. (2002) Detecting failure of BI : Monitoring information search in sequential bargaining. *Journal of Economic Theory*, 104, pp. 16-47.

Kagel, J., and Ross, A. (1995). Introduction. In *Handbook of Experimental Economics*, John Kagel and Alvin Roth (eds). Princeton University Press.

Keller, E. (2002) *Making Sense of Life*. Harvard University Press.

Khalil. E. (2000) Types of metaphors and identificational slips in economic discourse. *Research in the History of Economic Thought and Methodology: A Research Annual*, 18, pp. 83-105.

Klamer, H., & Leonard, T. (1994) So what's an economic metaphor? In *Natural Images in Economics*, Philipe Mirowski (ed). Cambridge University Press.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982) Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27, pp. 253-279.

Kuhn, H. (1953) Extensive games and the problems of information. In H. W. Kuhn and A. W. Tucker (eds). *Contribution of the Theory of Games*. Princeton University Press.

Kuhn, T. (1979 [1993]) Metaphor in science. In *Metaphor and Thought*, Andrew Ortony (ed). Cambridge University Press.

Lakoff, G. (1993) Contemporary theory of metaphor. In *Metaphor and Thought*, Andrew Ortony (ed). Cambridge University Press.

Lakoff, G., and Johnson, M. (1980) *Metaphors We Live By*. Chicago: Chicago University Press.

Leary, D. (1990) *Metaphors in the History of Psychology*. David Leary (ed). Cambridge University Press.

Lakoff, G., and Núnez, R. (1996) The metaphorical structure of mathematics: Sketching the cognitive foundations for a mind-based mathematics. In *Mathematical Reasoning: Analogies, Metaphors, and Images*, L English (ed). Hillsdale, NJ: Erlbaum.

Luce, D., and Raiffa, H. (1957) *Games and Decisions. Introduction and Critical Survey*. John Wiley and Sons, New York.

Mäki, U. (2002) The dismal queen of the social sciences. In *Fact and Fiction in Economics*, Uskali Mäki (ed). Cambridge, Cambridge University Press.

Maasen, S., Mendelsohn, E., and Weingart, P. (1995) *Biology as Society, Society as Biology: Metaphors*. Kluwer Academic Publishers.

McCloskey, D. (1983) The rhetoric of economics. *Journal of Economic Literature*, XXI, pp. 481-517.

McCloskey, D. (1990) Storytelling in economics. In *Economics and Hermeneutics*, Don Lavoie (ed). Routledge.

McCloskey, D. (1994) *Knowledge & Persuasion in Economics*. Cambridge University Press.

McKelvey, R., and Palfrey, T. (1992) An experimental study of the centipede game. *Econometrica*, 60, pp. 803-836.

Mincer, J. (1958) Investment in human capital and personal income distribution. *Journal of Political Economy*, 66, pp. 281-302.

Mirowski, P. (1988) Rhetoric, mathematics, and neoclassical theory. In *The Consequences of Economic Rhetoric*, A. Klamer, D. McCloskey, and R. Solow (eds). Cambridge University Press.

Mirowski, P. (1994) Doing what comes naturally: Four metanarratives on what metaphors are for. In *Natural Images in Economic Thought*, Philip Mirowski (ed). Cambridge University Press.

Morgan, M. (1999) Learning from models. In *Models as Mediators*, Mary Morgan and Margaret Morrison (eds). Cambridge University Press.

Morgan, M. (2002) Models, stories, and the economic world. In *Fact and Fiction in Economics*, Uskali Mäki (ed). Cambridge, Cambridge University Press.

Morgan, M. (2007) The curious case of the prisoner dilemma: Model situation? Exemplary narrative? In *Science Without Laws*, A. Creager, E. Lunbeck, and M. Norton Wise (eds). Durham: Duke University Press.

Morrison, M., and Morgan, M. (1999) Introduction. In *Models as Mediators*, Mary Morgan and Margaret Morrison (eds). Cambridge University Press.

Pettit, P., and Sugden, R. (1989) The BI paradox. *The Journal of Philosophy*, 86, pp. 169-182.

Pulaczewska, H. (1999) *Aspects of Metaphors in Physics. Examples and Case Studies*. Walter de Gruyter.

Rabinowicz, W. (1998) Grappling with the centipede game: Defence of BI for BI-terminating games. *Economics and Philosophy*, 14, pp. 95-126.

Richards, A. (1936) *The Philosophy of Rhetoric*. Oxford: Oxford University Press.

Reny, P. (1993) Common belief and the theory of games with perfect information. *Journal of Economic Theory*, 59, 257-274.

Rosenthal, R. (1981) Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25, pp. 92-100.

Roth, A., Vesna, P., Masahiro, O-F., and Shmuel, Z. (1991) Bargaining and market behavior in Jerusalem, Lubljana, Pittsburgh, and Tokyo: An experimental study. *American Economic Review*, 81, pp. 1068–1095.

Rubinstein, A. (2006) Dilemmas of an economic theorist. *Econometrica*, 74, pp. 865-883.

Schultz, T. (1961) Investment in human capital. *The American Economic Review*, 51, pp. 1-17.

Selten, R. (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4, pp. 25-55.

Selten, R. (1978) The chain store paradox. *Theory and Decision*, 9, pp. 127-159.

Sobel, J. (1993) Backward-induction arguments: A paradox regained. *Philosophy of Science*, 60, pp. 114-133.

Spiegler, P. (2006) Why quantitative social science needs interpretive methods. *Qualitative Methods*, 4, pp. 2-8.

Stalnaker, R. (1996) Knowledge, belief and counterfactual reasoning. *Economics and Philosophy*, 12, pp. 133-163.

Sugden, R. (1992) Inductive reasoning in repeated games. In *Rational Interaction*, Reinhard Selten (ed). Springer.

Sugden, R. (2002) The status of theoretical models in theoretical economics. In *Fact and Fiction in Economics*, Uskali Mäki (ed). Cambridge, Cambridge University Press.

Watson, J. (2013) *Strategy: An Introduction to Game Theory*. W. W. Norton & Company.