Computational Logic Seminar, Graduate Center CUNY

Beyond Nash

Sergei Artemov

December 1, 2009

In this talk

The knowledge-based rational decision model (*KBR*-model) suggests following a strategy yielding the highest payoff which the agent can secure to the best of his knowledge. Special (extreme) cases of *KBR* in perfect information (PI) games are the backward induction solution (assumes common knowledge of rationality), and the pure maximin solution (assumes ignorance of each other rationality).

In this talk, we prove a conjecture by A. Brandenburger that in PI games each *KBR*-path is a Nash path. Therefore, Nash equilibria capture *KBR*-solutions for all epistemic states of rational players, but cannot distinguish between them.

Game-theoretical assumptions

Our model of rational decision making uses standard game-theoretical assumptions, e.g., Harsanyi's Maximin Postulate ([6]),

> *If you cannot rationally expect more than your maximin payoff, always use a maximin strategy,*

and the traditional postulate of rational decision-making:

> *A rational player chooses a strategy that yields the highest payoff to the best of his knowledge.*

As noted in [1, 2], if a rational player operates in a non-probabilistic setting and bases his decision on **knowledge** rather than luck, guesswork, sudden opponent cooperation or error, etc., the aforementioned postulates lead to the same mathematical model of decision making that we call the Knowledge-Based Rational decision model (KBR-model).

Belief vs. Knowledge

Though Game Theory often considers decisions based on beliefs rather then knowledge (cf. [4]), a special theory of knowledge-based decision making looks to be appropriate as well. The principal difference between knowledge and belief is the factivity property of knowledge that beliefs do not necessarily possess. In some situations, players seem to make decisions on the basis of their knowledge and not merely on their beliefs: military, high-stakes commercial, juridical decisions, etc. Furthermore, according to commonly accepted properties of knowledge such as positive and negative introspection[1] ([5]), the decision-maker is aware of what he knows and what he does not know and hence is capable of distinguishing what he actually knows from what he merely believes without actual knowledge.

Highest Known Payoff vs Maximin

KBR suggests following a strategy that yields the highest payoff the agent can secure to the best of his knowledge. Equivalently, within the KBR approach, a rational player chooses a maximin solution over all strategies of others the player deems possible. These two seemingly different approaches produce the same result: a maximin choice over the set of all strategies a player considers possible (i.e., that cannot be ruled out as impossible) is a strategy yielding the highest guaranteed payoff to the best of that player's knowledge.

> Indeed, let $\mathbf{m}$ be the maximin payoff at a given node v over the set of all strategies that a player i considers possible. Then i knows a strategy that guarantees him payoff $\mathbf{m}$. On the other hand, for any other payoff $p > \mathbf{m}$, i knows that there is no strategy by i that could guarantee him payoff p. Therefore, $\mathbf{m}$ is the highest payoff that i knows he has a strategy for getting and he cannot rationally expect a payoff greater than $\mathbf{m}$.

Rationality Postulates

1. A rational player chooses a strategy yielding the highest payoff the agent can secure to the best of his knowledge. Equivalently, a rational player chooses a maximin solution over all strategy profiles the player deems possible.

2. Postulate (1) is commonly known and accepted by rational players.

Postulate (1) is the epistemically explicit form of Harsanyi's Maximin Postulate. Similarly, (2) is merely Harsanyi's Mutually Expected Rationality Postulate ([6]) expressed in epistemic language.

Strategies, profiles, paths

In this paper, we consider generic extensive-form perfect information games which include specification of the relevant states of knowledge for each player. In particular, for each player i, it is specified *which strategy profiles* σ *are known to be impossible* by player i. All other profiles are called *epistemically possible for player* i. By the factivity property of knowledge, no player is playing a strategy known to be impossible by any of the players.

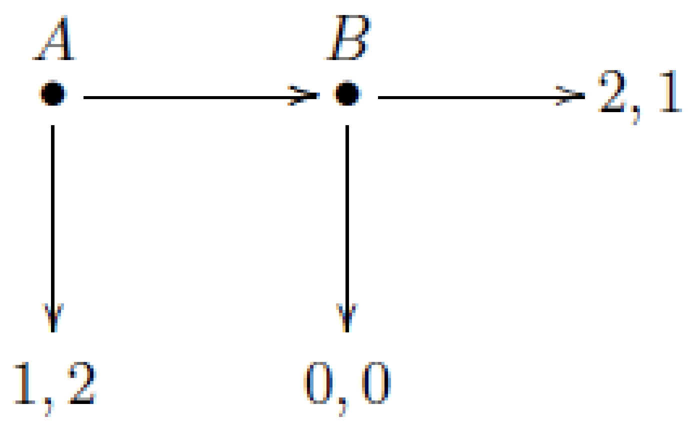
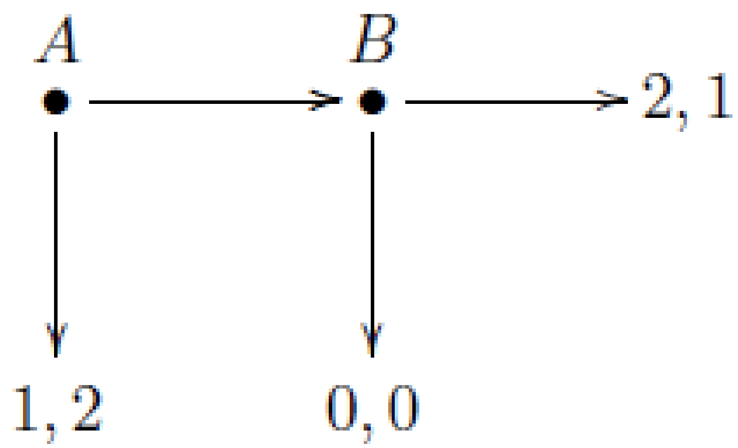


Figure 1: Game One, rationality is commonly known.

the only strategy which is epistemically possible is

$$\{across_A, across_B\}$$

which happens to also be the backward induction solution.



rationality is not mutually known.

each player considers each strategy of the other player possible. B knows that he is playing $across_B$. The epistemically possible strategy profiles for B are

$$\{down_A, across_B\}, \quad \{across_A, across_B\}.$$

Player A considers either strategy by B possible and cannot rationally expect to get a payoff greater than 1 if he plays $across_A$. By Rationality Postulate 1, A cannot choose $across_A$. Therefore, for A, the epistemically possible strategy profiles are

$$\{down_A, down_B\}, \quad \{down_A, across_B\}.$$

Note that though there is more than one strategy profile epistemically possible for each player, A and B each have a unique, epistemically possible strategy, namely $down_A$ for A and $across_B$ for B; we call them *KBR-strategies*. the *KBR*-path is $down_A$.

Subgames

For each node v of game G, a *subgame* G_v is determined by the rooted subtree with root v: epistemically possible strategy profiles for i in G_v are epistemically possible strategy profiles for i in game G relativized to the nodes from the subtree with root v.

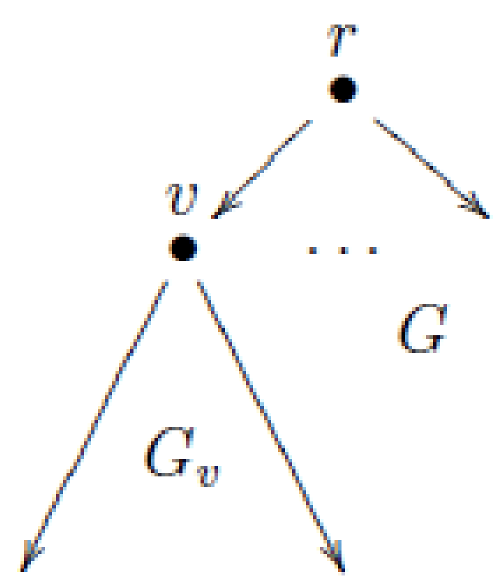


Figure 3: Subgame G_v of game G.

KBR strategy profile

Lemma 1 [1, 2] *At each node of a generic perfect information game, there is a unique move (called a KBR-move) by the corresponding player that yields the highest payoff that player can secure to the best of his knowledge (called highest known payoff).*

Corollary 1 *In a generic game with rational players, there is a unique KBR-move at each node.*

Definition 1 *A KBR-stategy for a given player i is a collection of KBR-moves at nodes where i makes a move.*

Corollary 2 *For each generic game with rational players, there is a unique KBR strategy profile and players actually play this profile.*

KBR view of the game

These observations lead to the following informal picture of epistemically possible strategy profiles for each rational player A; here B is any player other than A. At a node at which A makes a move, only the KBR-move is

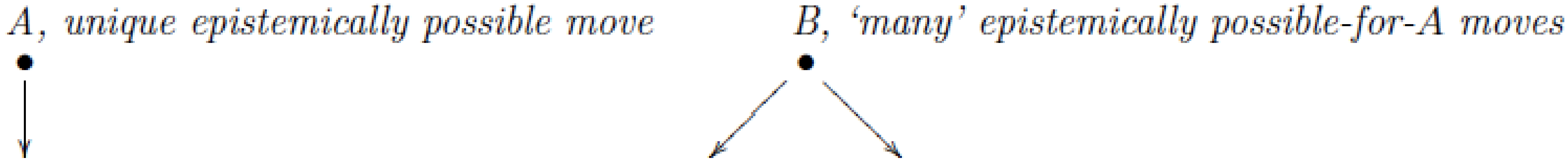


Figure 4: Strategy profiles that (rational) player A considers possible.

epistemically possible for A. At a node at which some other player makes a move, A may consider multiple moves as epistemically possible. All epistemically possible strategy profiles for A are constituted from A's unique KBR-strategy σ_A and strategies by others considered epistemically possible by A.

Corollary 3 [1, 2] *The real payoff for each player at a given node is greater than or equal to the highest known payoff at this node.*

BI and Maximin are special cases

Pure maximin strategy for a given player i corresponds to the reading of a game in which i has no information whatsoever about other players' epistemic states. Then i considers all moves by opponents epistemically possible. Under these conditions, pure maximin is a special case of KBR.

Another special case of the KBR-solution is given by the backward induction solution BI under Aumann's conditions of common knowledge of rationality ([3]). In this case, each player has sufficient information to exactly determine his opponents' move at each node. For each player, there is only one epistemically possible strategy profile: the KBR-solution of the game.

Each *KBR*-path is Nash

KBR strategy profile is not necessarily a Nash profile.

$$
\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & 2,1 \\
1,2 & 0,0 &
\end{tikzcd}
$$

rationality is not mutually known.

The KBR strategy profile is

$$\sigma = \{down_A, across_B\},$$

and the KBR-path is

$$P = down_A.$$

σ is not a Nash profile.

$\sigma' = \{across_A, across_B\}$ yields A's payoff 2.

On the other hand, there is a Nash strategy profile

$$\sigma'' = \{down_A, down_B\}$$

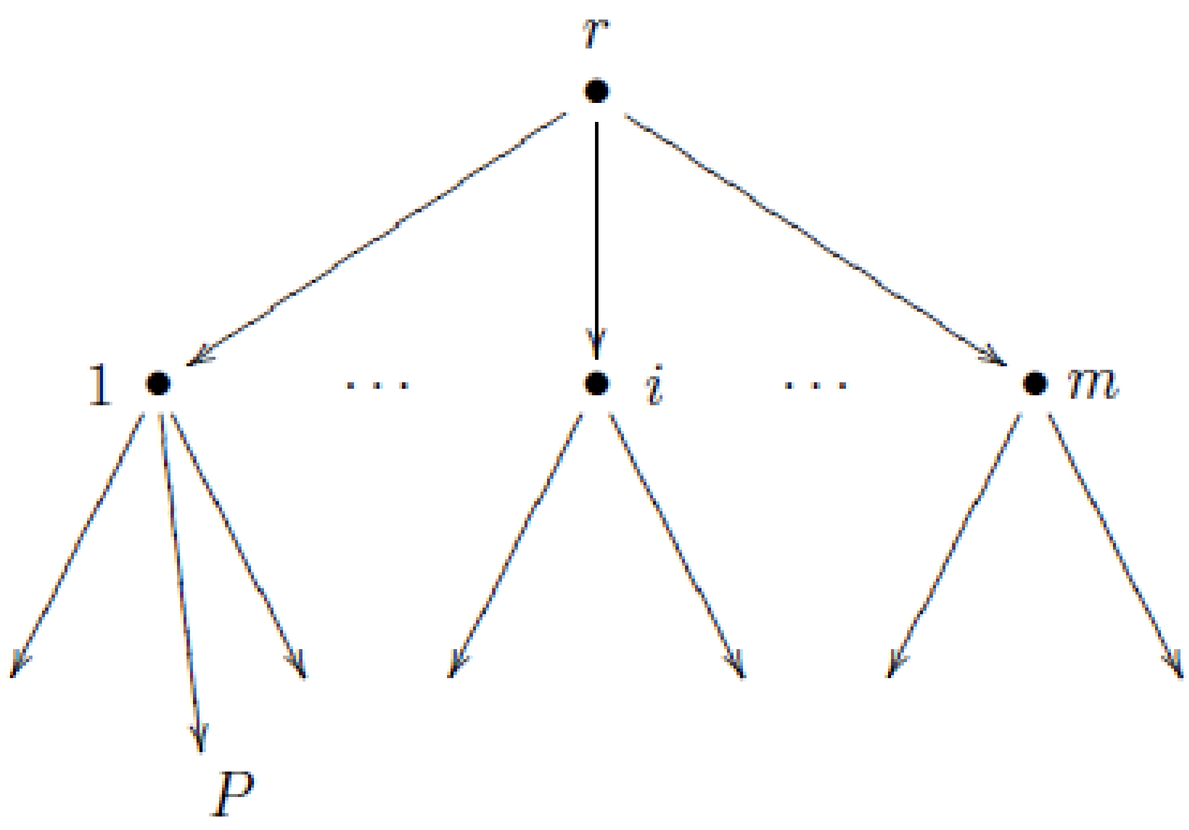
that has the same path P as σ.

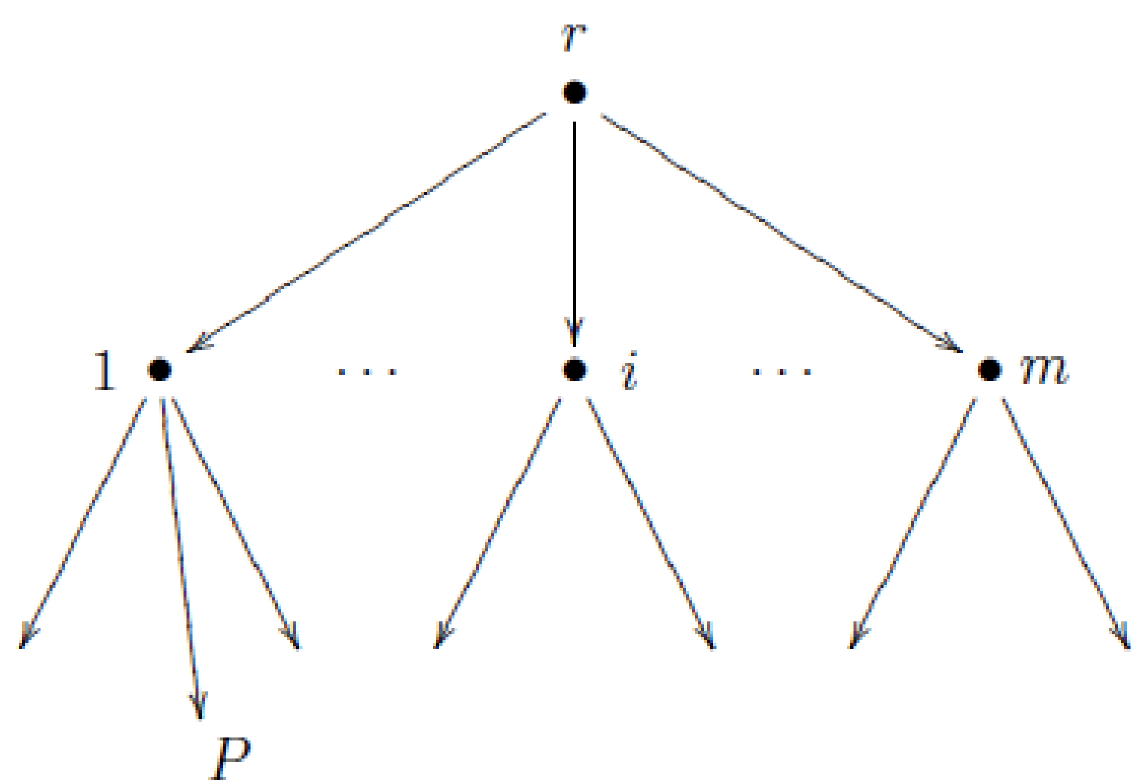
Theorem 1 *In a PI game with rational players, the KBR-path is a Nash path.*

Proof. Induction on maximal game length $n(G)$. The base: $n(G)$=1. Then the KBR-path consists of one rational move which constitutes a Nash profile.

The Induction Hypothesis: suppose the theorem claim holds for all games with length less than k.

The Induction Step. Consider a PI game G in an extensive tree-like form such that $n(G) = k$. Let P be its KBR-path, A be the player who is making a move at root node r, and $1, \ldots, m$ be immediate successors to r. By $G_1, \ldots, G_m$, we denote subgames of G with roots at $1, \ldots, m$ respectively.



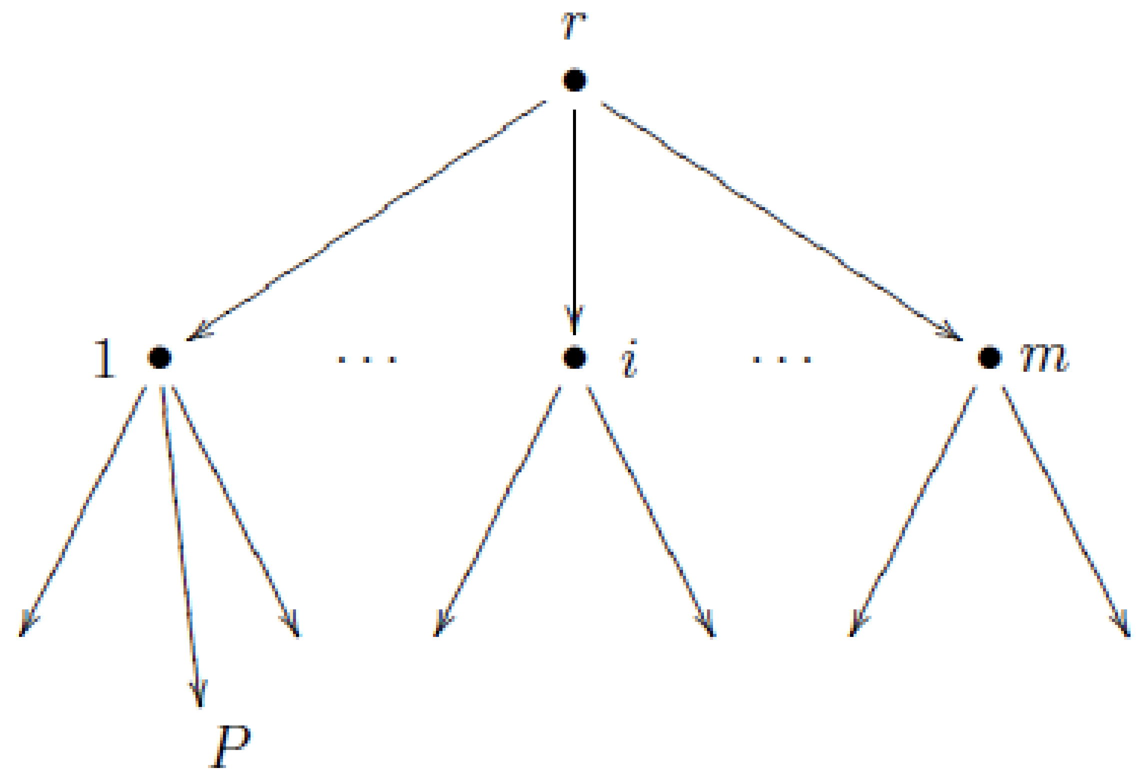


Let b be the highest known payoff for player A at root node r (cf. [1]), i.e., the highest payoff that A knows he can secure at r:

$$b = \mathbf{HKP}_A(r).$$

Then for any strategy σ_A by A, there is a strategy profile σ containing σ_A and epistemically possible for A such that A's payoff of σ, $U_A(\sigma)$ is less than or equal to b. By Corollary 3, A's payoff on path P, $U_A(P)$ is greater than or equal to b.

Without loss of generality, assume that A's root move is $(r, 1)$, and that the rest of P, P_1 occurs within G_1. By Lemma 2, P_1 is the *KBR*-path in G_1. By the Induction Hypothesis, since $n(G_1) < k$, P_1 is a Nash path in G_1, i.e., there is a Nash strategy profile σ^1 such that P_1 is its path in G_1.



Our goal now is to extend σ^1 to a Nash strategy profile σ for all of G without changing its path P. For this, we have to define the moves of each player at nodes other than those from G_1.

At root node r, A's move is $(r, 1)$ as suggested by P: make it part of σ. It now remains to define moves at all nodes of games $G_2, \ldots, G_m$.

Pick subgame G_i, $i = 2, \ldots, m$ and consider the following auxiliary 'maximin game' on the same tree. In this maximin game, player A tries to win more than his highest known payoff b, and all other players are playing against this goal. Label a leaf **S** (for **Success**) if A's payoff at this leaf is greater than b, and **F** (for **Failure**) otherwise. Backward induct to label all other nodes of G_i and define moves for each node v of G_i.

Case 1. A makes a move at node v, and all immediate successors to v are labeled **F**. Then label v as **F** and pick an arbitrary move for A at v.

Case 2. A makes a move at node v, and there is an immediate successor to v that is labeled **S**. Then label v as **S** and pick a move for A from v to one of its immediate **S**-successors.

Case 3. A player other than A makes a move at v, and there is an immediate successor to v labeled **F**. Then label v as **F** and pick a move from v to one of its immediate **F**-successors.

Case 4. A player other than A is making a move at v, and all immediate successors to v are labeled **S**. Then label v as **S** and pick an arbitrary move at v.

Let us denote σ_A^i the strategy by A, and σ^i_{-A} the collection of strategies by all other players in the maximin game on G_i.

The following lemma shows that A cannot win the maximin game.

Lemma 3 *The root node i of G_i is labeled* **F**.

Lemma 3 *The root node i of G_i is labeled* **F**.

Proof. Since b is the highest known payoff for A at the root node, given σ^i_A, there should be a collection δ^i_{-A} of strategies for other players in G_i (deemed possible by A) such that A's payoff of the profile $\{\sigma^i_A, \delta^i_{-A}\}$ is less than or equal to b. Let P' be the path of $\{\sigma^i_A, \delta^i_{-A}\}$ in G_i. We claim that each node of P' is labeled **F**. Backward induction on the length of P'. The leaf of P' is labeled **F** since it indicates A's payoff on P' which is not greater than b. Consider a node v of P' whose immediate successor in P' is labeled **F**. If v is an A-node[3], all immediate successors to v in G_i are labeled **F**, hence v is labeled **F**. If v is a non-A-node[4], v is labeled **F** as well. So all nodes of P' are labeled **F**, including the root node i of G_i. $\square$

Now we define the desired strategy profile σ on G_i-nodes:

For each $i = 2, \ldots, m$, σ restricted to G_i-nodes coincides with $\{\sigma^i_A, \sigma^i_{-A}\}$.

σ's path is P since the first move of σ is $(r, 1)$
and the rest of the path is P_1.

Lemma 4 *σ is a Nash strategy profile.*

Proof. Present σ as a collection of A's strategy σ_A and non-A-strategies σ_{-A}.

Players other than A cannot improve their payoff by unilaterally deviating from σ_{-A} given σ_A. Indeed, changes outside G_1 do not alter the outcome. Changes inside G_1 cannot improve the payoff since within G_1, σ is a Nash strategy profile.

Fix σ_{-A} and consider an arbitrary strategy σ'_A for A.

Case 1. The first move of σ'_A is $(r, 1)$. Then the consequences of σ'_A are limited to changes in A's strategy within G_1 that cannot yield a better payoff for A, since σ is a Nash profile on G_1.

Case 2. The first move of σ'_A is (r, i) with some $i = 2, \ldots, m$. Suppose, en route to contradiction, that

$$U(\{\sigma'_A, \sigma_{-A}\}) = b' > b$$

Lemma 4 *σ is a Nash strategy profile.*

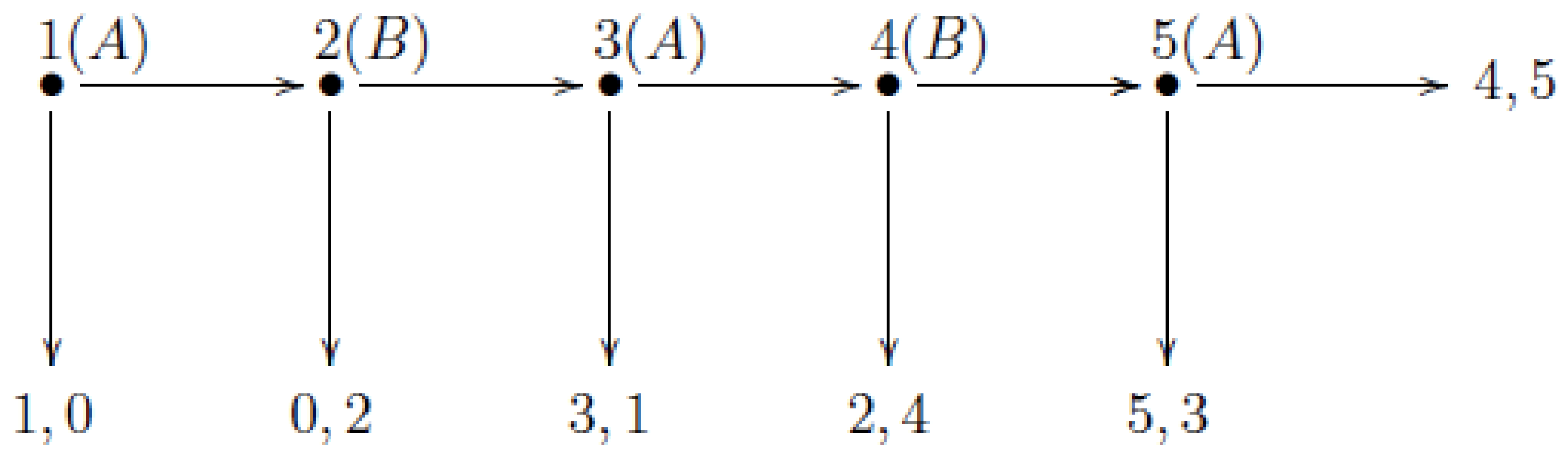
Case 2. The first move of σ'_A is (r, i) with some $i = 2, \ldots, m$. Suppose, en route to contradiction, that

$$U(\{\sigma'_A, \sigma_{-A}\}) = b' > b$$

and let P' be the path in G_i corresponding to $\{\sigma'_A, \sigma_{-A}\}$. By backward induction on the node depth, we show that all nodes of P' are labeled **S**. Base: the leaf node of P' is labeled **S** since P' delivers A's payoff $b' > b$. Let v be a node in P' whose immediate successor in P' is labeled **S**. If v is an A-node, then v should be labeled **S** by definition of the labeling process. If v is a non-A-node, then P''s move at v is made according to σ, which indicates that all immediate successors to v in G_i are labeled **S**, hence v is labeled **S** as well.

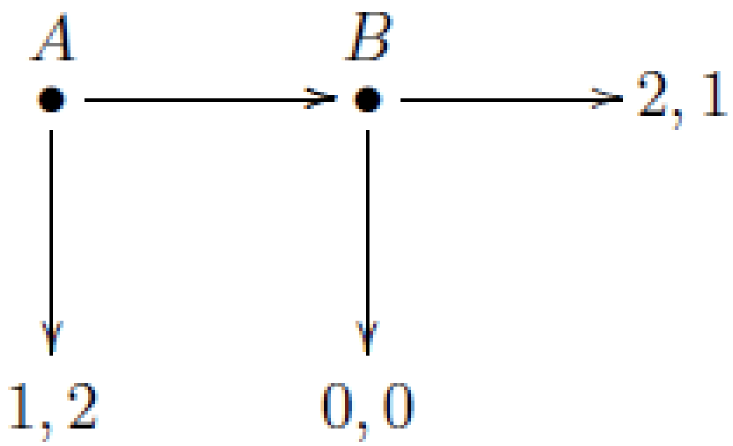
We have arrived at a contradiction to Lemma 3 that states i is labeled **F**. This proves Lemma 4. $\square$

As an easy corollary to this theorem, we conclude that each backward induction path is a Nash path. Indeed, apply Theorem 1 to the variant of the game in which common knowledge of rationality is assumed. For such games, the resulting *BI*-path will be the *KBR*-path. By Theorem 1, this path is Nash. Likewise, each pure maximin path is Nash as well, since the maximin profile is the *KBR*-profile with players ignorant of each others' rationality.



Corollary. *In the centipede game under all epistemic states of players the solution path is* ***down*** *at the first node.*

Proof. There is a unique Nash path.



rationality is not mutually known.

None of the strategies is strictly dominating, hence the *IU* strategy profiles here are

$$\{down_A, down_B\},\ \{across_A, down_B\},\ \{down_A, across_B\},\ \{across_A, across_B\}.$$

NE strategy profiles

$$\{down_A, down_B\} \text{ and } \{across_A, across_B\},$$

and two *NE*-paths

$$down_A \text{ and } (across_A, across_B).$$

one *KBR*-path that happens to be the *MAXM*-path as well: $down_A$.

There is one *BI*-path $(across_A, across_B)$,
that is also the *KBR*-path in a version of Game One in which common knowledge of rationality of players is assumed.

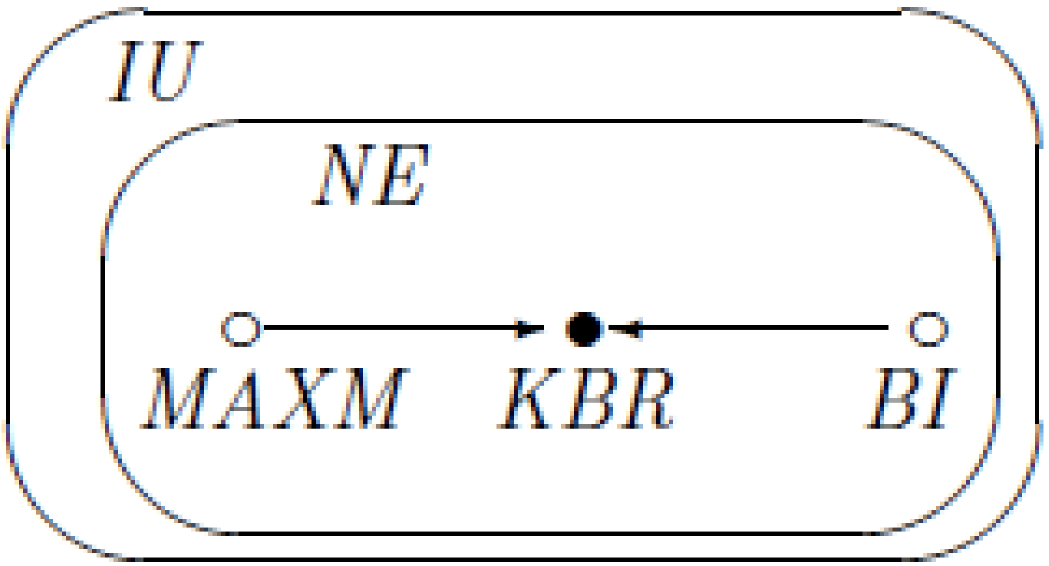


Figure 7: Comparing methods

The empty bullets are a reminder that *MAXM*- and *BI*-paths are justified only under special conditions, e.g., complete ignorance of each others' rationality (*MAXM*), or common knowledge of rationality (*BI*). The arrows indicate that under corresponding conditions, the *MAXM*-path and *BI*-path become the *KBR*-path.

KBR vs Aumann's rationality

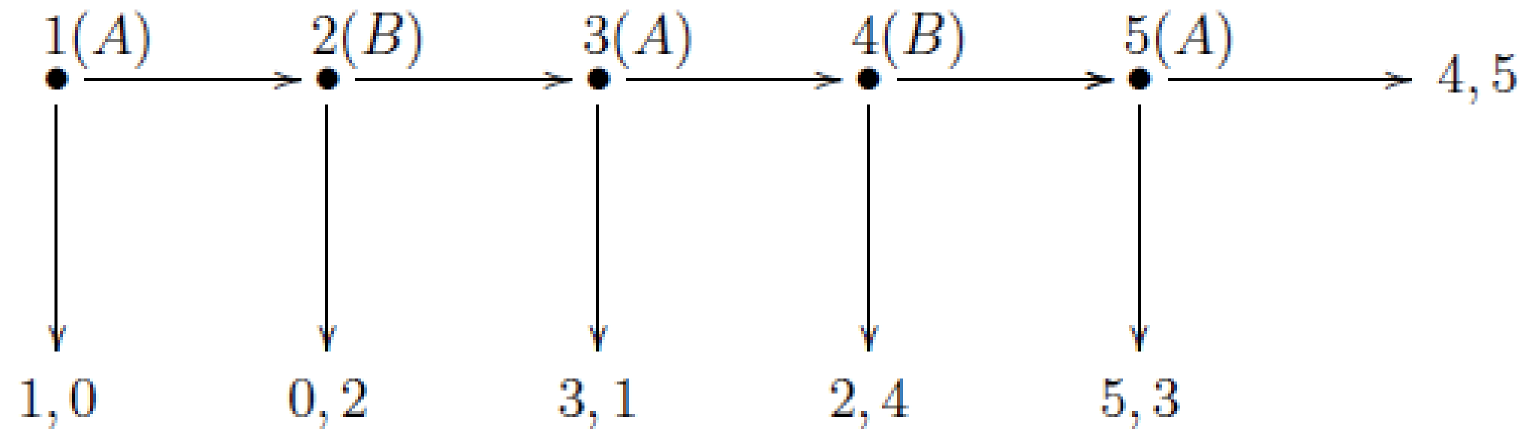
The mathematical formulation of Aumann's rationality considers *irrational* any choice by i of a strategy σ_i in a situation when i **knows** that there is another strategy σ_i' which strictly dominates σ_i. All strategies which are not irrational are considered *rational*.

This definition of rationality works well when uncertainty at a node can be completely resolved by epistemic reasoning, e.g., under common knowledge of rationality in perfect information games, or when the player knows that one strategy strictly dominates all others. However, Aumann's rationality does not help to make decisions in general situations, e.g., when there is a choice of several strategies, none of which strictly dominates the others.

Any KBR-rational strategy is Aumann-rational, but not vice versa.

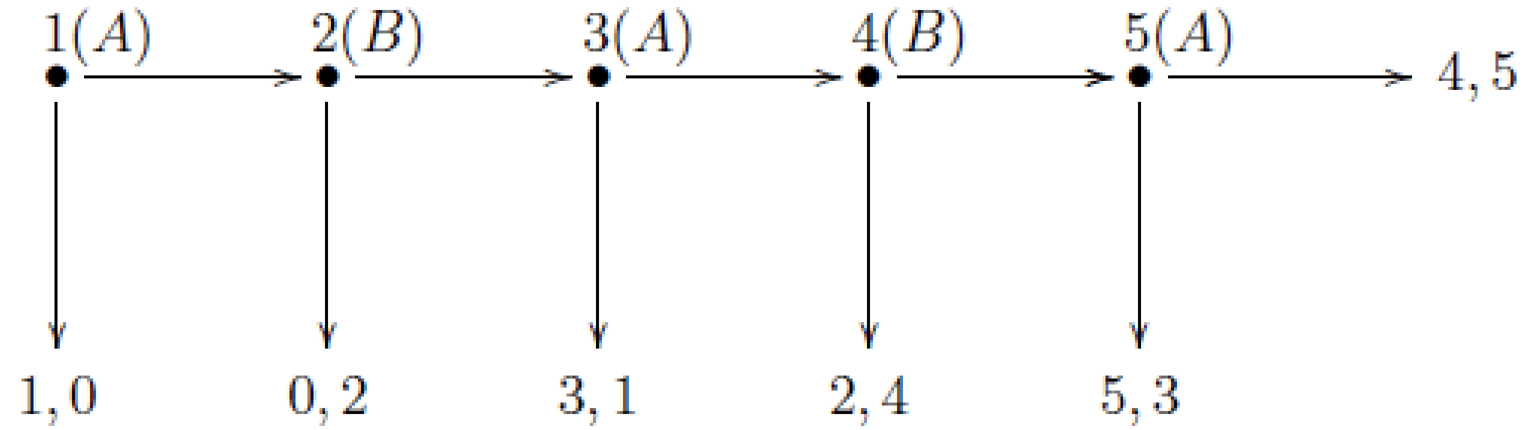
Finally, if Aumann's rationality yields a definitive answer, then this answer is the *KBR*-solution. In this respect, *KBR*-rationality may be regarded as a definitive extension of Aumann's rationality.

The power of public announcement PA



Game Five. Consider a version of the Centipede game in Figure 2 in which A decides to make one irrational move across at node 5, and this, along with rationality of players at all other nodes, is commonly known. Then both players choose across at all nodes, A's payoff is 4, and B's payoff is 5. Indeed, by assumption, A plays across at 5. At node 4, B knows that A will play across at 5, hence B rationally plays across at 4 as well. B's reasoning is known to A, hence A plays across at 3, etc.

PA with a probabilistic twist



Game Seven. Player A can do even better if he is allowed to play probabilistically. Suppose, in Figure 2, at node 5, A plays across with probability 0.51 and down with probability 0.49, and this, along with the rationality of players at all other nodes, is commonly known. Then B's expected payoff when playing across at node 4 is 4.02 which is a higher payoff than playing down at node 4. So B rationally chooses across at 4. Likewise, A's expected payoff at node 3 when playing across is 4.49, which is higher than his payoff of down at node 3. Therefore, A plays across at 3. The same reasoning justifies playing across at nodes 1 and 2 as well. As a result, both players play across at nodes 1–4, and A tosses a 0.51/0.49 coin at node 5. The average payoff for A is 4.49 (and can be made arbitrary close to 4.5 by playing with 'almost' 0.5 probability at node 5) and the average payoff for B is 4.02.

Modernization Dilemma and PA

Two players: **Gov** - government, **Corp** - corporation.
Corp faces an overdue modernization. Moves: *modernize* and *stagnate*.
Gov wants to take over. Moves: *free* and *control*.

	free		*control*
modernize	**2,2**	$\Leftarrow$	**0,0**
	$\Downarrow$		$\Downarrow$
stagnate	**3,1**	$\Rightarrow$	**1,3**

Stagnate is the dominant strategy for **Corp**.
The only Nash equilibrium is (*stagnate*, *control*) with payoffs (**1**,**3**).
Corp realizes this and plays strategically: publicly announcing *modernize* !
By elimination of possible worlds rule, the second row disappears. The new equilibrium is (*modernize*, *free*) with payoffs (**2**,**2**).

Acknowledgments

To all Computational Logic Seminar participants.

Special thanks to Adam Brandenburger and Amanda Friedenberg whose paper "Self-Admissible Sets" was the source of the Brandenburger Conjecture. They discovered that epistemically motivated game paths (in their version) should all be Nash paths, and this my work just confirms their observation for *KBR*.

My thanks also go to Vladimir Krupski, Elena Nogina, and Cagil Tasdemir for useful comments and corrections.

Many thanks to Karen Kletter for editing this paper.