Computational Logic Seminar, Graduate Center CUNY

Rational decisions in non-probabilistic setting

Sergei Artemov

October 20, 2009

In this talk

The knowledge-based rational decision model (*KBR*-model) offers an approach to rational decision making in a non-probabilistic setting, e.g., in perfect information games with deterministic payoffs. The *KBR*-model is an epistemically explicit form of standard game-theoretical assumptions, e.g., Harsanyi's Maximin Postulate. This model suggests following maximin strategy over all scenarios which the agent considers possible to the best of his knowledge.

In this talk, we compare *KBR* with other approaches and show that *KBR* is the only non-probabilistic decision making method which is definitive, rational, and based exclusively on knowledge.

Uncertainty without probabilities?

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

Suppose A is mission control, and has the option of sending into space a specially trained astronaut B who, unfortunately, has been exposed to German measles or a reserve astronaut (payoff 1). If B does not get sick, his mission will be a success (payoff 2), otherwise it will be aborted with failure (payoff 0), cf. Figure 1.

Uncertainty without probabilities?

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

With enough good will, we can apply **Harsanyi's Maximin Postulate** ([12], sections 6.2 and 6.3, postulate A1) here:

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

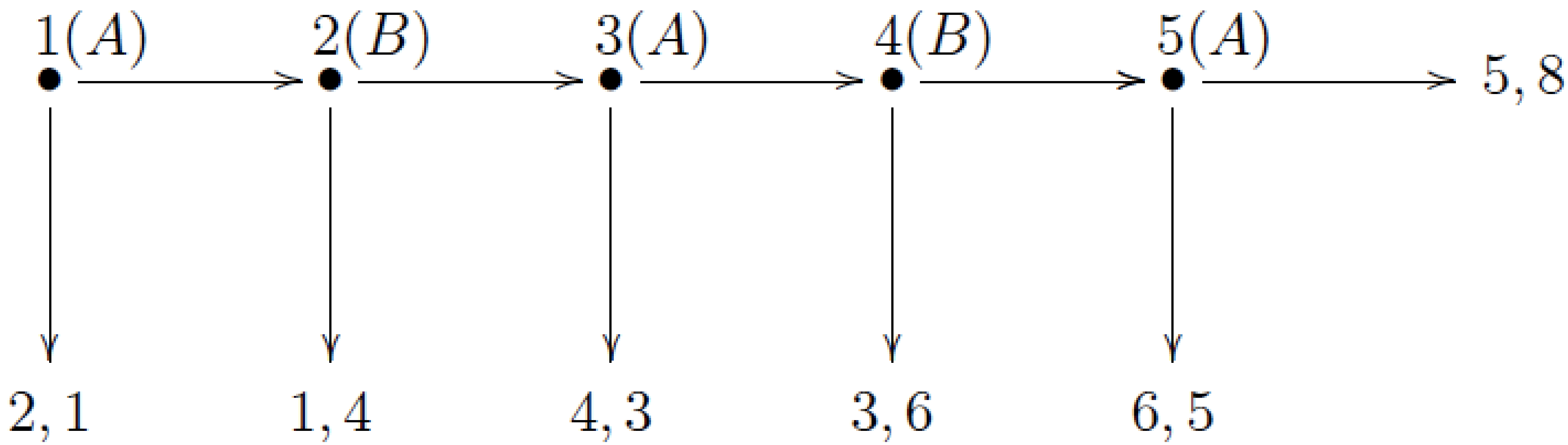
According to our scenario, A can hope, but cannot know for sure, that B does not get sick and delivers payoff 2. Therefore, A has no reason to "rationally expect" more than maximin value 1 when moving *across*, so the rational choice for A is the maximin solution *down*[1].

Epistemic Game Theory

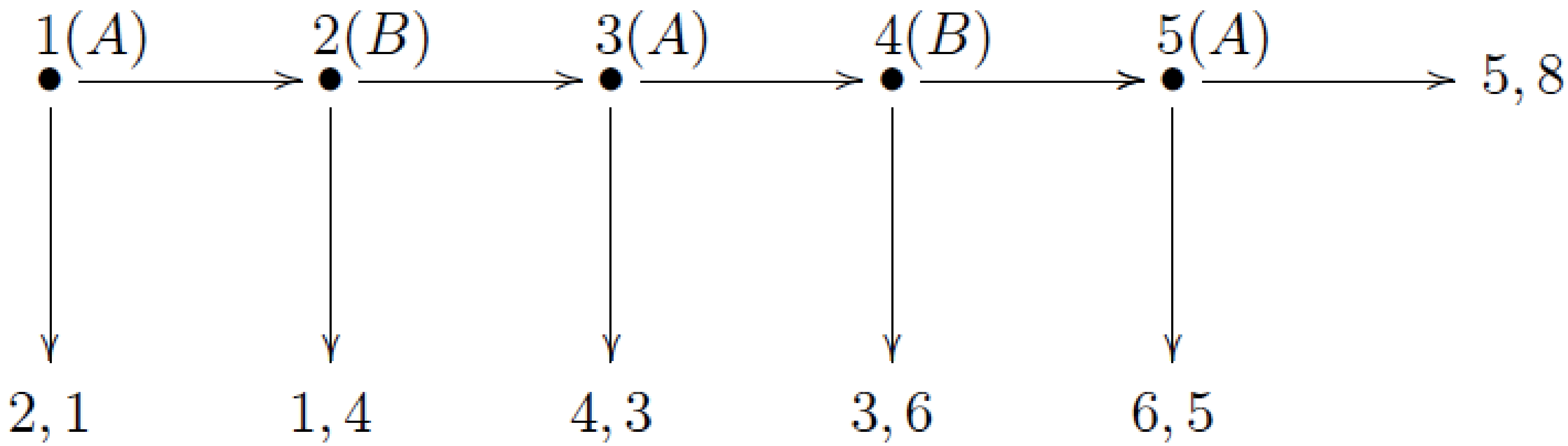
First admitted that epistemic states matter and studied conditions under which standard game theoretical solutions hold (backward induction, Nash, etc.).

Is still on the way towards developing a coherent theory of games in which epistemic states of players are a legitimate part of the game specification?

Centipede



The standard *backward induction* solution of Centipede is offered by Aumann's Theorem on Rationality ([3]) which assumes *common knowledge of rationality* and predicts playing *down* at each node. Indeed, at node 5, player A's rational choice is *down*. Player B is certainly aware of this and, anticipating A's rationally playing *down* at 5, would himself play *down* at 4. Player A understands this too, and would opt *down* at 3, etc. This analysis has been used in textbooks and expository articles as an illustration of Aumann's Theorem (cf. [18]).



Theorem 1 [1] *In the Centipede game, under any states of players' knowledge, rational players play down at each node.*

Proof. At node 5, player A choses *down*.

At node 4, player B's maximin strategy is to play *down*. In addition, B **cannot know** that A would play *across* because A in fact plays *down*[4]. Therefore, B cannot rationally expect to get more than his maximin payoff of 6 at node 4 **regardless of his knowledge about** A. By Harsanyi's Maximin Postulate, B chooses maximin strategy *down*.

At node 3, player A cannot not know that B will play *across* at 4, since B in fact plays *down*. Therefore, A cannot rationally expect to exceed his maximin payoff of 4 and hence, by Harsanyi's Maximin Postulate, plays *down*.

Likewise, B plays *down* at 2 and A plays *down* at 1. □

Another paradigm: knowledge

A naïve, pre-epistemic understanding of rational decisions is

A rational player chooses a strategy which yields the highest payoff. (1)

This formulation captures the 'greedy' element of rationality, i.e., going for the highest payoff, but totally ignores its epistemic component: as we have already agreed, a rational player decides not on the basis of what is true in the world, but rather on the basis of what he knows/believes. In particular, the highest actual payoff associated with a given strategy can be unknown to the player, who therefore will not be able to take this payoff into account. The knowledge requirement naturally leads to the following epistemically explicit reading of (1):

A rational player chooses a strategy which yields the highest **known** *payoff.*

Maximin and Knowledge converge

Rationality Postulates:

I. *A rational player in perfect information games chooses a maximin solution among all strategies the player deems possible.*

II. *Postulate* (I) *is commonly known and accepted by rational players.*

Postulate I is the epistemically explicit form of Harsanyi's Maximin Postulate. Likewise, (II) is nothing but Harsanyi's Mutually Expected Rationality Postulate expressed in epistemic language.

In the rest of this section, we will show that

the maximin solution among all strategies the player deems possible

corresponds to

the best known payoff to the best of player's knowledge.

Strategies, moves, outcomes...

A *strategy* of player i is a function that assigned an *action* (a.k.a. *move*) to each node of the game in which i is making decision. A strategy profile

$$\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$$

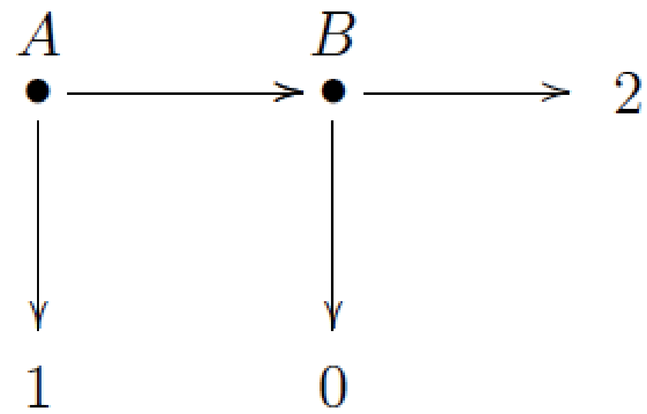
as a collection of strategies σ_i for players $i = 1, 2, \ldots, n$. Each strategy profile σ uniquely determines the *outcome* $O(v, \sigma)$ associated to σ and a node v: $O(v, \sigma)$ is the *terminal history* (i.e., the sequence of moves by players from v to a terminal node) in which each move is made according to σ. Likewise, everyone who knows the game tree can calculate i's payoff determined by σ and a node v: $U_i(v, \sigma)$.

Highest Known Payoff of a strategy

The following observation first made in [1][5], is the foundation of *KBR*-theory:

Proposition 1 *Let v be a node and i the player who makes a move at v*[6]*. Then for every strategy σ_i of i, there is a unique Highest Known Payoff, $HKP_i(v, \sigma_i)$, equal to the minimum of i's payoffs for all strategy profiles σ containing σ_i which are deemed possible by i at v:*

$$HKP_i(v, \sigma_i) = \min\{U_i(v, \sigma) \mid \sigma \textit{ is a possible strategy profile containing } \sigma_i\}.$$



For example, in the decision making schema in Figure 1, A has two strategies: to play $down_A$ with payoff 1 or to play $across_A$ and pass the choice to B whose intentions are unknown to A. The game tree, including payoffs, is supposed to be known to A. Under these conditions, for agent A,

$$HKP_A(A, down_A) = 1,$$

whereas

$$HKP_A(A, across_A) = 0,$$

since A considers both strategies by B, $across_B$ and $down_B$ possible, hence the strategy profile

$$\{across_A, down_B\}$$

is possible for A and brings A's payoff 0. Given strategy $across_A$, out of two payoffs, 0 and 2, A knows that he gets at least 0, but does not know whether he gets payoff 2.

The situation is different in Game Two in Figure 3. B is now assumed to be a rational player who has his own payoffs (coinciding with those of A). Suppose also that A knows that B is rational.

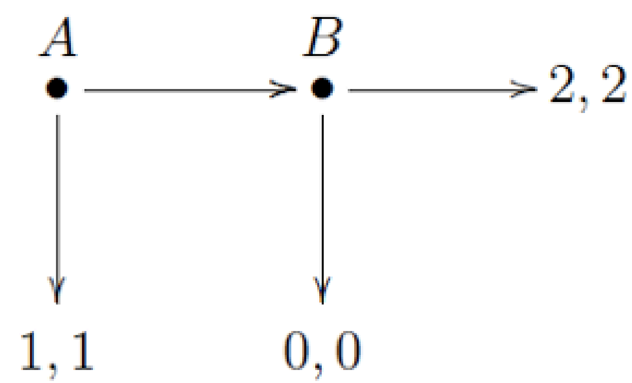


Figure 3: Game Two

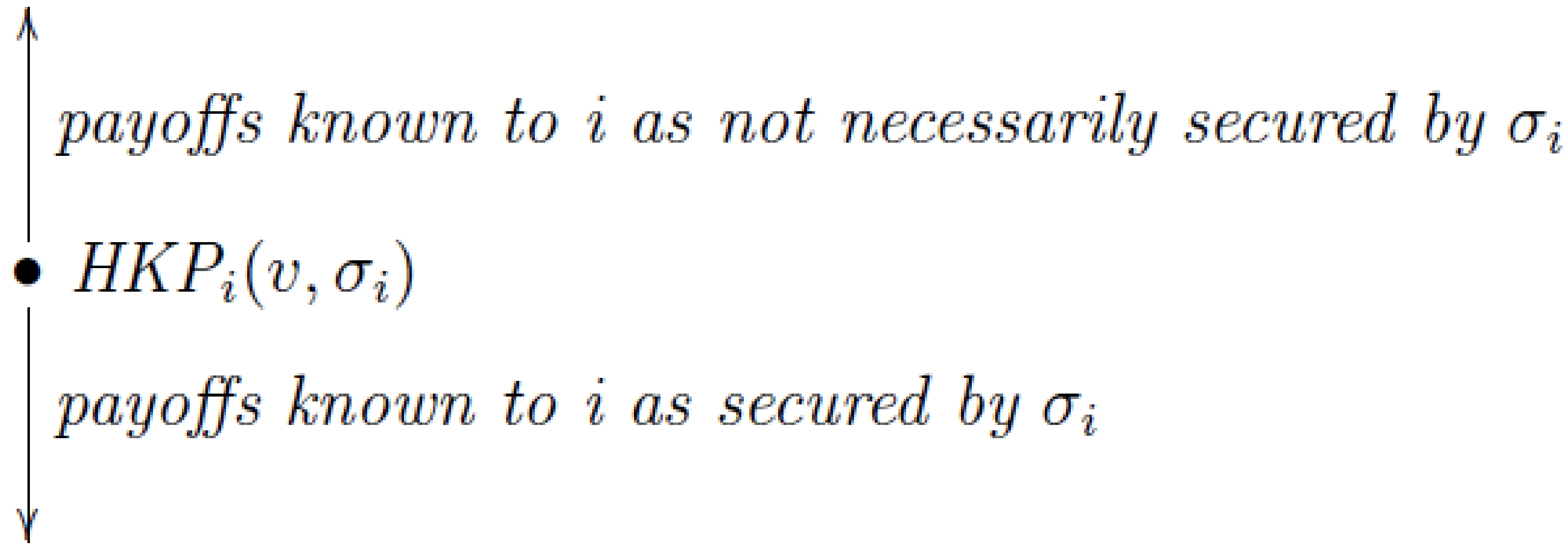
Under these assumptions, B would play across seeking payoff 2 for himself (and for A as well). Moreover, this is now known to A! Therefore, A knows that B will not play down, hence for A the only possible strategy profile containing $across_A$ is

$$\{across_A, across_B\},$$

with A's payoff 2. Under these circumstances,

$$HKP_A(A, across_A) = 2.$$

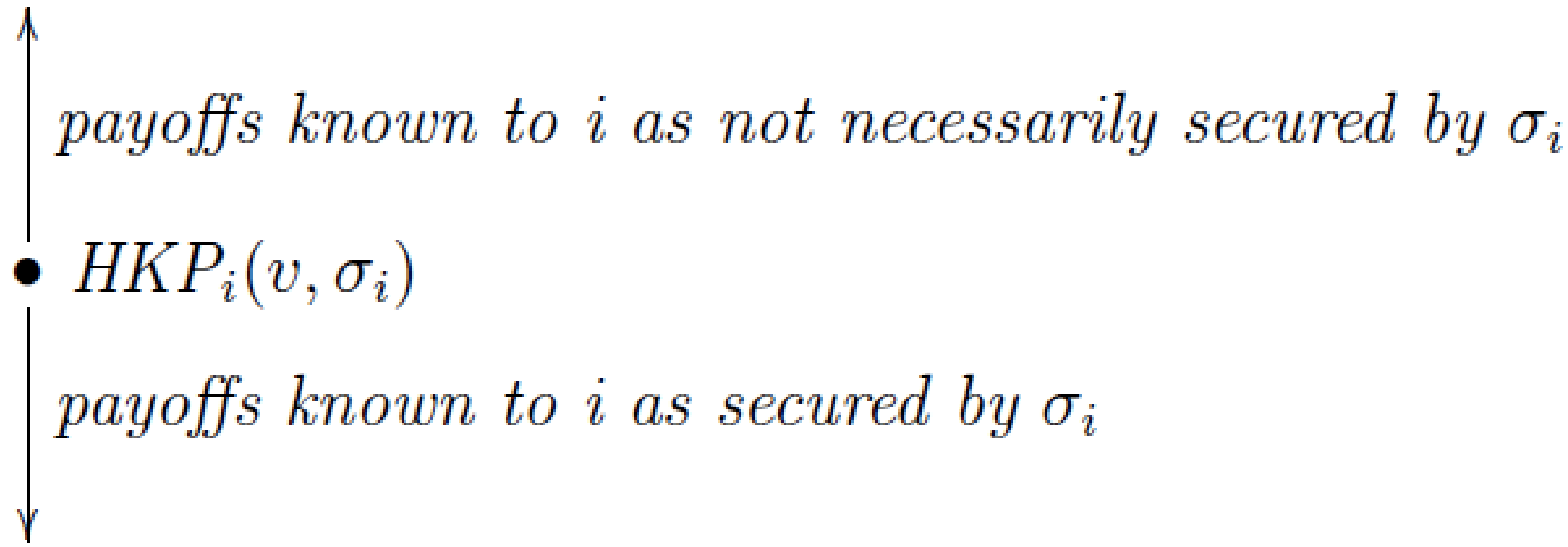
In a version of Game Two where A is not aware of B's rationality, A considers both moves by B possible and $HKP_A(A, across_A) = 0$ again.



The epistemic role of the Highest Known Payoff of a strategy can be illustrated by Figure 4. For each node v, player i makes and strategy σ_i of i, the (finite) set of all possible i's payoffs is naturally divided into two intervals.

Interval 1: $\{p \mid p \leq HKP_i(v, \sigma_i)\}$. For each payoff p from this interval, player i knows that he will get at least p when playing σ_i from v .

Interval 2: $\{q \mid q > HKP_i(v, \sigma_i)\}$. For each payoff q from this interval, player i knows that if he plays σ_i, then his opponents have response strategies which i cannot rule out at the best of his knowledge and which bring i a payoff strictly less than q. So payoff q is known to i as not necessarily secured by strategy σ_i.



How constructive is this knowledge? Player i considers some of the strategy profiles possible, and some not, at the best of his knowledge. For each strategy σ_i, and each possible strategy profile σ containing σ_i, i can calculate i's payoff $U_i(v, \sigma)$ of σ starting from v. Since $HKP_i(v, \sigma_i)$ is the minimum of $U_i(v, \sigma)$ over all such σ containing σ_i, $HKP_i(v, \sigma_i)$ is known to player i.

Highest Known Payoff of a move

Definition 1 *Highest Known Payoff of a move $\mathcal{M}$ at a node v, $HKP_i(v, \mathcal{M})$, is the maximum of Highest Known Payoffs of strategies that start with move $\mathcal{M}$ at v:*

$$HKP_i(v, \mathcal{M}) = \max\{HKP_i(v, \sigma_i) \mid \textit{ for all strategies } \sigma_i \textit{ making move } \mathcal{M} \textit{ at } v\}.$$

It is clear that $HKP_i(v, \mathcal{M})$ is known to i and attainable when i plays the strategy σ_i which realizes the maximum of $HKP_i(v, \sigma_i)$'s.

Best Known Strategy

Definition 2 *Strategy* σ_i *is a* **best known strategy** *for player i at a given node v if* σ_i *has the greatest highest known payoff among all i's strategies, i.e.,*

$$HKP_i(v, \sigma_i') \leq HKP_i(v, \sigma_i)$$

for all i's strategies σ_i'.

Best Known Move

A game is called **generic** if there are no indistinguishable payoffs for each player.

The best known strategy is not necessarily unique even for generic games, e.g., if strategies differ at some node which is not accessible during the game, then these strategies are formally different but apparently yield the same payoff under each response strategy and hence have the same *HKP*. However, if we limit our attention to the **first move** of a strategy, then we come to the notion of *the best known move* for a given player at a given node, which is already unique for generic games. This notion reflects the idea of a definitive rational choice.

Definition 3 *A move $\mathcal{M}$ for player i at a node v is the* **best known move** *if it is the first move of a best known strategy for i at v.*

Equivalently, the best known move is that which has the greatest *HKP*, cf. Definition 1 and [1].

Best Known Move: uniqueness

Theorem 2 [1] *At each node v of a generic game, there exists a unique best known move at v.*

Proof. Let i be the player who makes a move at v. Existence follows from the fact that for each strategy σ_i, there is a well-defined highest known payoff $HKP_i(v, \sigma_i)$. To prove uniqueness, note that best known strategies are those which have the greatest $HKP_i(v, \sigma)$, which is, of course, unique for i at a node v, by definition. We claim that all best known strategies at a given node of a generic game start with the same move. Indeed, if two strategies start with different moves in the game tree, they have disjoint sets of terminal nodes and hence disjoint sets of payoffs for a given player. Such strategies could not have the same HKP. $\square$

Maximin meets Knowledge

Note that the best known strategy is a clear maximin notion projected on the epistemic state of the player. We first determine the highest known payoff for each strategy σ_i by taking the minimum payoffs for all possible strategy profiles containing σ_i, then taking the maximum for all strategies σ_i.

Definition 4 *The Highest Known Payoff for player i at a node v, $HKP_i(v)$, is the maximum of $HKP_i(v, \sigma_i)$'s for all strategies σ_i by i:*

$$HKP_i(v) = \max\{HKP_i(v, \sigma_i) \mid \textit{for all strategies } \sigma_i \textit{ by } i\}.$$

Naturally, $HKP_i(v)$ equals $HKP_i(v, \sigma_i)$ for the best known strategy σ_i at v.

Figure 5 shows that $HKP_i(v)$ separates two intervals:

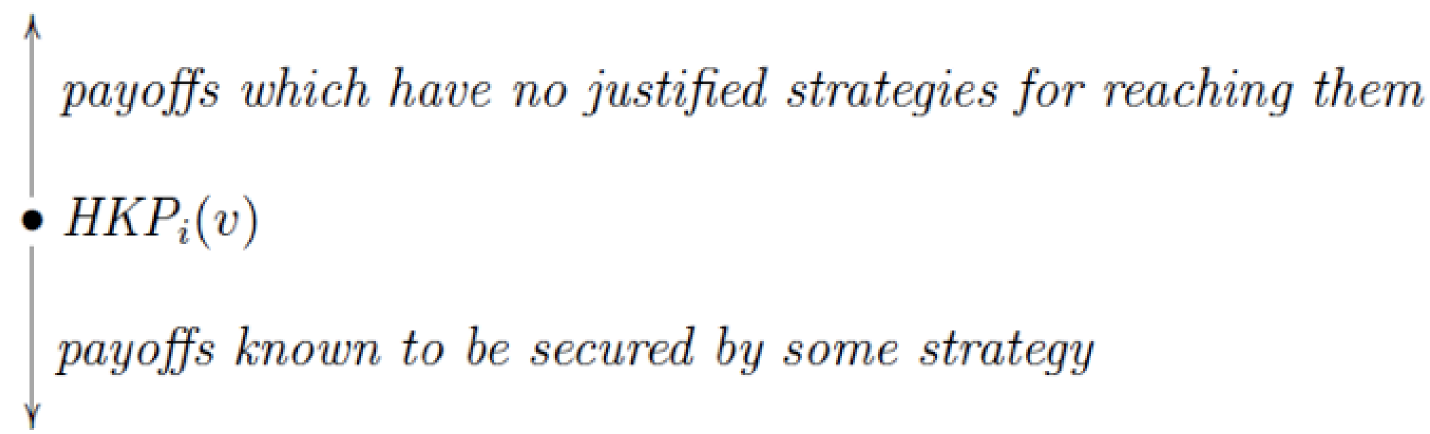


Figure 5: Highest Known Payoff at a node

Interval 1: $\{p \mid p \leq HKP_i(v)\}$. These payoffs are known to i to be secured by a certain strategy, and i knows exactly which strategy to use to secure these payoffs.

Interval 2: $\{q \mid q > HKP_i(v)\}$. Neither of the payoffs from this region has a justified strategy for player i to secure this payoff. Unless the game description says otherwise, i cannot rationally expect more than the maximin payoff $HKP_i(v)$ at node v.

KBR decision method

A *decision method* for a certain class of games is a prescription of choosing moves at the nodes of the game. There are several well-known decision methods used for perfect information games: Nash equilibrium and subgame perfect equilibrium, backward induction solution, pure maximin, (iterated) eliminating of dominated strategies, etc.

Definition 5 *Knowledge-based rational decision (KBR) method chooses a move with the best knows payoff at each node.*

Why KBR is so special for PI games?

The goal of Section 2 has been to argue that in perfect information games with their finite game trees and deterministic payoffs and no other assumptions for making a decision, *KBR* decision method of choosing a move with the best knows payoff at each node is consistent with the epistemically correct form of Harsanyi's Maximin Postulate, Rationality Postulate I, In this Section we compare *KBR* with other decision methods and show that *KBR* is the only definitive method consistent with the epistemic form of Harsanyi's Maximin Postulate.

A decision method is *knowledge-based rational* if its choice of action (move) is consistent with Rationality Postulates I and II.

A decision method is *definitive* for a certain class of games if it provides a unique choice of action (move) at each node of every generic game from this class. This condition rules out speculative 'solutions' such as 'all moves are rational,' etc.

Why KBR is so special for PI games?

Theorem 3 [KBR Theorem] *For perfect information games, the knowledge-based rational decisions is the only decision method which is definitive and knowledge-based rational.*

Before we proceed with proving this theorem, consider some examples of decision methods.

1. Nash equilibrium and subgame perfect equilibrium;
2. Backward induction solution;
3. Pure maximin;
4. Eliminating dominated strategies.

Nash and subgame perfect equilibria

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

There are two Nash equilibria

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

which allow A to move either way, hence this method is not definitive.

Neither of these equilibria is uniformly rational. Indeed, if A knows that B is rational, A's rational choice is across since A knows that B will then play across and deliver payoff of 2 to A. Therefore, the second of the equilibria, $\{down_A, down_B\}$, is not rational for A. If A does not know that B is rational, then A considers both moves by B possible and, by Harsanyi's Maximin Postulate, A should rationally play *down*. In this case, equilibrium $\{across_A, across_B\}$ is not rational for A.

Subgame perfect equilibrium (cf. [14]) when applied to Game Three eliminates equilibrium $\{down_A, down_B\}$, which is the only rational solution for A if A considers both moves by B possible. Hence this method is not necessarily rational.

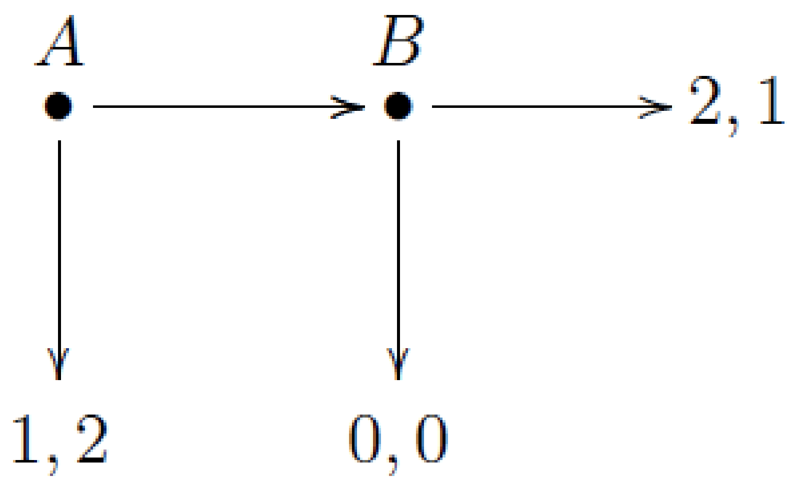
Backward induction

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

The backward induction solution is rational and knowledge-based, but not definitive. Moreover, backward induction is a method of avoiding uncertainty by calculating opponents' strategies. In particular, backward induction does not provide answers under uncertainty.

In Game Three, if A knows that B is rational, then the backward induction solution is $\{across_A, across_B\}$. However, if B is rational, but A does not know this and considers both moves by B possible, then backward induction does not work, thus leaving A without any recommendation at all.

Pure Maximin



Pure maximin is not necessarily rational. In Game Three, if A knows that B is rational, then A's best strategy is $across_A$ and this strategy is known by A to bring him a payoff of 2. The pure maximin solution $down_A$ brings A payoff 1 and hence is not rational.

Eliminating dominated strategies

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

Eliminating dominated strategies is an epistemically correct method which, however, is not definitive. We refer the reader to [14] for exact definitions and restrict our attention here to an example.

In Game Three, where A considers both moves by B possible, neither of A's strategies, $across_A$ and $down_A$, is dominated and hence cannot be eliminated. Therefore, eliminating dominated strategies alone does not provide a definitive answer here.

Proof of KBR-theorem

Proof. Fix a generic game and a node v. The *KBR*-method consists of choosing the move which yields the highest known payoff for given player i at v, $HKP_i(v)$. Such a move exists and is unique (Theorem 2).

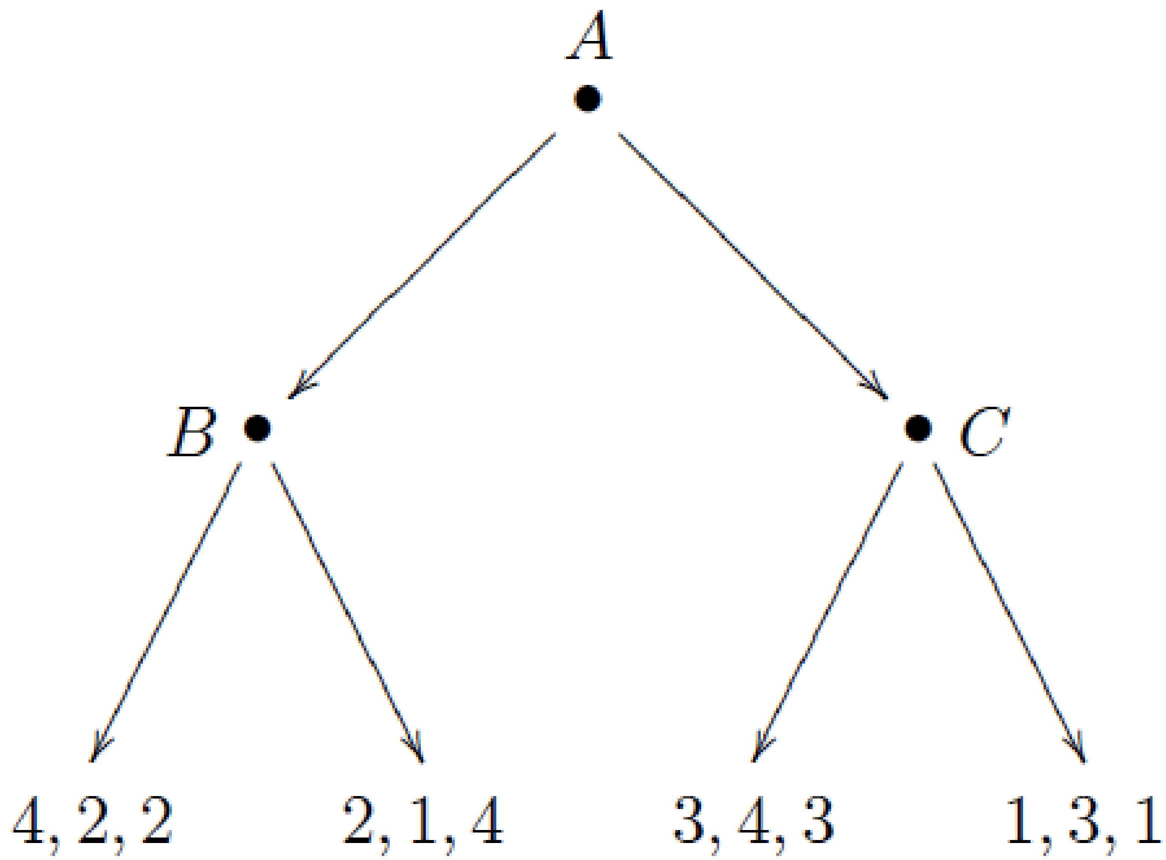
Suppose i has to make a move at v and chooses move $\mathcal{M}$. Let m be the highest payoff which i knows is secured by $\mathcal{M}$. Since $\mathcal{M}$ is definitive, by Theorem 2, there is a unique such m.

Case 1. $m = HKP_i(v)$. Then $\mathcal{M}$ is the *KBR* choice since for generic games, by Theorem 2, different moves have different *HKP*'s.

Case 2. $m < HKP_i(v)$. Choosing $\mathcal{M}$ contradicts the Harsanyi's Maximin Postulate - Rationality Postulate I since $\mathcal{M}$ is not the maximin move for i at v. Indeed, the maximin move is the best known move at v corresponding to $HKP_i(v)$, which is different from m.

Case 3. $m > HKP_i(v)$. This case is impossible, by definition of $HKP_i(v)$. Indeed, the highest payoff m which i knows is secured by move $\mathcal{M}$ cannot be higher than the highest known payoff at v. $\square$

Active manipulation



Suppose *A* is not aware of *B* and *C*'s rationality. Then *A* moves *left* to secure payoff 2. Actually, *A* gets 4 which is more than expected. Suppose also that *B* and *C* are smart enough to understand this. Then *B* can manipulate *A* by leaking the true information that ***C* is rational**. *A* then knows that *right* secures his payoff 3, which is higher than *A*'s known payoff of *left*: *A* plays *right* and gets 3 (less), *B* gets 4 (much more) and *C* gets 3 (more). *C* does not have an incentive to disclose that *B* is rational, hence

B wins without ever making a move!

Full knowledge is power

Model predictions:

- ***Every game with rational players has a solution.*** Rational players know which moves to make at each node.

- ***Those who know the game in full know its solution***, i.e., know everybody's moves.

Partial knowledge can hurt

Model predictions:

- More knowledge yields a ***higher known payoff*** but not necessarily a ***higher actual payoff***. So

  nothing but the truth

  can be misleading.

- Knowing

  the whole truth

  however, yields a higher actual payoff.

When knowledge does not matter

Model predictions:

In ***strictly competitive (e.g. zero-sum) games***, all players' epistemic states lead to the same (maximin) solution.

Maybe this is why military actions (typical zero-sum games) do not require sophisticated reasoning about other players:

just do it

normally suffices.

Conclusions

Do we recommend playing perfect information games using KBR strategy?

1. Not if you can responsibly assign probabilities to your opponents' responses, otherwise

2. To the best of your knowledge, rule out all impossible strategies of the game. If some uncertainly remains, it's this: you cannot know more. Deal with this uncertainty using KBR; this is the only rational method of playing PI games.

Acknowledgments

To all Computational Logic Seminar participants who have had patience to listen to so many iterations of this work.

Special thanks to Adam Brandenburger, Mel Fitting, Vladimir Krupski, Loes Olde Loohuis, Elena Nogina, Graham Priest, and Cagil Tasdemir.

Many thanks to Karen Kletter for selflessly editing endless versions of this paper.