UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

When all is done but not (yet) said: Dynamic rationality in extensive games

Baltag, A.; Smets, S.; Zvesper, J.A.

Publication date
2008

Published in
Workshop on Logic and Intelligent Interaction, ESSLLI 2008

Link to publication

Citation for published version (APA):
Baltag, A., Smets, S., & Zvesper, J. A. (2008). When all is done but not (yet) said: Dynamic rationality in extensive games. In *Workshop on Logic and Intelligent Interaction, ESSLLI 2008* (pp. 58-73) http://ai.stanford.edu/~epacuit/LaII/proceedings/zvesper.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:20 Apr 2025

When all is done but not (yet) said: Dynamic rationality in extensive games

Alexandru Baltag
Oxford
baltag@comlab.ox.ac.uk

Sonja Smets
Brussels & Oxford
sonsmets@vub.ac.be

Jonathan A. Zvesper
Amsterdam
jonathan@illc.uva.nl

The jury is still out concerning the epistemic conditions for backward induction, the "oldest idea in game theory" ([2, p. 635]). Aumann [2] and Stalnaker [31] take contradictory positions in the debate: Aumann claims that common 'knowledge' of 'rationality' in a game of perfect information entails the backward-induction solution; Stalnaker that it does not.[1] Of course there is nothing wrong with any of their relevant formal proofs, but rather, as pointed out by Halpern [22], there are differences between their interpretations of the notions of knowledge, belief, strategy and rationality. Moreover, as pointed out by Binmore [14, 15], Bonanno [17], Bicchieri [13], Reny [26], Brandenburger [18] and others, the reasoning underlying the backward induction method seems to give rise to a *fundamental paradox* (the so-called "BI paradox"): in order to even start the reasoning, a player assumes that (common knowledge, or some form of common belief in) Rationality holds at all the last decision nodes (and so the obviously irrational leaves are eliminated); but then, in the next reasoning step (going backward along the tree), some of these (last) decision nodes are eliminated, as being incompatible with (common belief in) Rationality! Hence, the assumption behind the previous reasoning step is now undermined: the reasoning player can now see, that *if* those decision nodes that are now declared "irrational" were ever to be reached, then the only way that this could happen is if (common belief in) Rationality failed. Hence, she was *wrong* to assume (common belief in) Rationality when she was reasoning about the choices made at those last decision nodes. This whole line of arguing seems to undermine itself!

In this paper we use as a foundation the relatively standard and well-understood setting of Conditional Doxastic Logic (CDL, [16, 5, 7, 6]), and its "dynamic" version (obtained by adding to CDL operators for *truthful public announcements* $[!\varphi]\psi$): the logic PAL-CDL, introduced by Johan van Benthem [11]. In fact, we consider a slight extension of this last setting, namely the logic APAL-CDL, obtained by further adding dynamic operators for *arbitrary* announcements $[!]\psi$, as in [3]). We use this formalism to capture a novel notion of "dynamic rationality" and to investigate its role in decision problems and games. As usual in these discussions, we take a *deterministic stance*, assuming that the initial state of the world at the beginning of the game already fully determines the future play, and thus the unique outcome, irrespective of the players' (lack of) knowledge of future moves. We do *not*, however, require that the state of the world determines what *would* happen, if that state were *not* the actual state. That is, *we do not need to postulate the existence of any "objective counterfactuals"*. But instead, we only need "subjective counterfactuals": in the initial state, not only the future of the play is specified, but also the players' *beliefs* about each other, as well as their *conditional beliefs*, pre-encoding their *possible revisions of belief*. The players' conditional beliefs express what one may call their "propensities", or "dispositions", to revise their beliefs in particular ways, if given some particular pieces of new information.

Thus at the outset of a game, all is "done", including the future. But all is not necessarily said. In a deterministic model, as time progresses the only thing that changes are the pictures of the world in the minds of the players: the information states of the players. This is *"on-line" learning*: *while* the game is being played, the players *learn* the played moves, and so they may change their minds about the situation. We can *simulate* this on-line learning (and its effect on the players' beliefs) via *off-line "public announcements"*: if, *before* the start of the game, the agents were *publicly told* that the game will reach some node u, then they would be in the *same epistemic state* as they would have been by (not having any such public announcement but instead) playing the game until node u was reached.

So in this paper we stress the importance of the *dynamics* of beliefs and rationality during a play of an extensive game, and we use dynamic operators in order to simulate the play of the game. Since we focus on games of *perfect information*, we only need *public* announcements to simu-

[1] Others agree with Stalnaker in disagreeing with Aumann: for example, Samet [29] and Reny [26] also put forwards arguments against Aumann's epistemic characterisation of subgame-perfect equilibrium. Section 5 is devoted to a discussion of related literature.

late the moves of the game. The idea of adding modalities for public announcements to epistemic logic was introduced and developed in [24, 20]. Dynamic epistemic logic [4] provides for much richer dynamic modalities than just public announcements, capturing the effects of more complex and more "private" forms of learning. We think these could be applied to the case of games with *imperfect information*. However, for simplicity, we leave these developments for future work and consider for now only perfect information, and so only public announcements.

Using the terminology of Brandenburger [18], ours is a *belief-based approach* to game theory (in the same category as the work of Battigalli and Siniscalchi [9, 10]), in contrast to the *knowledge-based approach* of Aumann [2] and others. This means that we take the players' *beliefs* (including conditional beliefs) as basic, instead of their *knowledge*. However, there is a notion of knowledge that naturally arises in this context: the "irrevocable knowledge", consisting of the beliefs that are *absolutely unrevisable*, i.e. believed under any conditions. This notion of knowledge is meant to apply only to the players' "hard information", obtained by observation or by undoubtable evidence. This is a much stronger condition than "certain belief" (subjective probability 1) or even "true belief", and as a result it may happen that very few things are "known" in this sense. One of the things we assume to be irrevocably known is the *structure of the game*: the possible outcomes, the players' preferences etc; also, in a game of perfect information, the *played moves* are observed, and thus known, after they are played; finally, another thing irrevocably known to a player is *her own beliefs*: by introspection, she knows what she believes and what not. Besides this, we do not assume much else to be known, although our setting is definitely consistent with (common) knowledge of all the players' beliefs, their strategies, their rationality etc.

One thing we do *not* assume as known is the *future of the game*: no outcomes that are consistent with the structure of the game are to be excluded at the outset of the game. In fact, we make the opposite assumption: that it is common knowledge that nobody knows the future, i.e. nobody knows that some outcome will not be reached. This "open future" assumption seems to contradict common knowledge of rationality; but in fact, it is consistent with it, if by rationality we only mean "rational planning", leaving open the possibility that players may make mistakes or may change their minds. The players may certainly *believe* their rational plans will be faithfully carried out, but they have no way to "know" this in advance. We think of our "open future" assumption as being a realistic one, and moreover one that embodies the agents' "freedom of choice", as well as the "possibility of error", that underly a correct notion of rationality. An agent's rationality can be assessed only if she is given some options to freely choose from. There are certainly cases in which the future can be known, e.g. when it is determined by a known natural law. But it is an essential feature of rational agents that their own choices are not known to them to be thus determined; or else, they would have no real choices, and thus no rational choice. Any natural determinism is assumed to be absorbed in the definition of the game structure, which does pose absolute limits to choices. In a sense, this simply makes precise the meaning of our "knowledge" as "hard information", and makes a strict delimitation between the past and the future choices, delimitation necessary to avoid the various paradoxes and vicious circles that plague the notions of rational decision and freedom of choice: the agents may have "hard information" about the past and the present, but not about their own future free choices (although they may have "soft" information, i.e. "certain" beliefs, with probability 1, about their future choices).

Our notion of *"dynamic" rationality* takes into account the dynamics of beliefs, as well as the dynamics of knowledge. On the one hand, following Stalnaker, Reny, Battigalli and Siniscalchi etc. (and in contrast with Aumann), we assess the rationality of a player's move at a node against the beliefs held *at the moment when the node is reached*. On the other hand, we incorporate the above-mentioned epistemic limitation to rationality: the rationality of an agent's move only makes sense when that move is *not already known* (in an irrevocable manner) to her. Agents cannot be held responsible for moves that they cannot choose or change any more. Since the agents' knowledge increases during a game of perfect information, their set of available options decreases: passed options/nodes, or nodes that were bypassed, cannot be the objects of choice any more. As a result, our notion of rationality is *future-oriented*: it only concerns her plans concerning current and future decisions. An agent can be rational *now* even if in the past she has made some "irrational" moves. So in a sense, the meaning of "rationality" *changes in time*, synchronous to the change of beliefs and the change of (known) set of options. This concept of rationality, developed on purely a priori grounds, solves in one move the "BI-paradox": the first reasoning step in the backward-induction argument (dealing with the last decision nodes of the game) is *not undermined* by the result of the second reasoning step, since the notion of "Rationality" assumed in the first step is *not* the same as the "Rationality" disproved in the second step! The second step only shows that some counterfactual nodes cannot be reached by rational play, and thus it implies that some agent must have been irrational (or must have had some doubts about the others' rationality, or must have made some "mistake") *before* such an "irrational" node was reached; but this doesn't contradict in any way the assumption that the agents will be rational *at* that node (and further in the future).

Since dynamic rationality is only about *rational plan-*

ning, we need to strengthen it in order to capture *rational playing* of the game. We do this by adding to dynamic rationality a condition requiring that players actually play in accordance with their beliefs. The resulting condition is called "rational play".

Dynamics cannot really be understood without its correlative: *invariance* under change. Certain truths, or beliefs, *stay true* when everything else changes. We have already encountered an "absolute" form of invariance: "irrevocable knowledge", i.e. belief that is invariant under *any* possible information change. Now, we need a second, weaker form of invariance: "stability". A truth, or a belief, is *stable* if it remains true, or continues to be believed, after any *(joint) learning of "hard" information* (via some truthful public announcement). In fact, in the case of an "ontic" (non-doxastic) fact p, Stalnaker's favourite notion of "knowledge" of p [31, 33] (a modal formalisation of Lehrer and Klein's "defeasibility theory of knowledge"), also called "safe belief" in [7], corresponds precisely to *stable belief* in p. Stability can be or not a property of a belief or a common belief: a proposition P is a "stable (common) belief" if the fact that P is (common) belief is a stable truth, i.e. P continues to be (common) belief after any (joint) learning of "hard" information.

We can now give an informal statement of the main theorem of this paper:

> *Common knowledge of the game structure, of "open future" and of stable (common[2]) belief in rational play entails common belief in the backward induction outcome.*

Overview of the Paper To formalise stability and "stable common belief", we introduce in the next section Conditional Doxastic Logic CDL and its dynamic version APAL-CDL. Section 2 recalls the definition of extensive games and shows how to build models of those games in which the structure of the game is common knowledge, in our strong sense of "knowledge". In Section 3 we define "rationality" and "rational play", starting from more general decision-theoretic considerations, and arriving at a definition of *dynamic rationality* in extensive (aka "dynamic") games, which is in some sense a special case of the more general notion. Section 4 gives a formal statement of our main results. Section 5 discusses connections between our work and some existing literature on the epistemic foundations of backward induction.

[2] Adding the word "common" to this condition doesn't make a difference: common knowledge that everybody has a stable belief in P is the same as common knowledge of common safe belief in P.

1 Conditional Doxastic Logic

CDL models, also called "plausibility models" are essentially the "belief revision structures" in Board [16], simplified by incorporating structurally the assumption of Full Introspection of Beliefs (which allows us to use binary plausibility relations on worlds for each agent, instead of ternary relations). But since we will also want to talk about the *actual change* under the effects of actions, like moves in a game, rather than just the *static* notion that is in effect captured by Board's models, we will enrich the language of CDL with model-changing *dynamic* operators for "public announcements", in the spirit of Dynamic Epistemic Logic (cf. [4, 11, 12]).

The models are "possible worlds" models, where the worlds will usually be called *states*. Grove [21] showed that the AGM postulates [1] for rational belief change are equivalent to the existence of a suitable pre-order over the state space.[3] The intended interpretation of the pre-order $\leq_i$ of some agent i is the following: $s \leq_i t$ means that, in the event $\{s, t\}$, i considers *s at least as plausible as t*.

In interactive situations, where there are several players, each player i has a doxastic pre-order $\leq_i$. In addition to having different *beliefs*, any two players might have different *knowledge*. We follow the mainstream in game theory since Aumann and model interactive knowledge using a partitional structure. However, as in Board [16], we will derive i's partition from i's pre-order $\leq_i$. Let us be more precise: fix a set S and a relation $\leq_i \subseteq S \times S$; then we define the *comparability class* of $s \in S$ for $\leq_i$ to be the set $[s]_i = \{t \in S \mid s \leq_i t \text{ or } t \leq_i s\}$ of states $\leq_i$-comparable to s. Now we want the set of comparability classes to form a partition of S, so we will define a *plausibility frame* to be a sequence $(S, \leq_i)_{i \in N}$ in which S is a non-empty set of states, and each $\leq_i$ a pre-order on S such that for each $s \in S$, the restriction of $\leq_i$ to $[s]_i$ is a "complete" (i.e. "total" or "connected") preorder.

Fact 1.1 *In any plausibility frame, $\{[s]_i \mid s \in S\}$ forms a partition of S. We will interpret this as the* information partition *for player i (in the sense of "hard" information, to be explained below).*

So we can define player i's *knowledge* operator in the standard way, putting for any "proposition" $P \subseteq S$:

$$K_i P := \{s \in S \mid [s]_i \subseteq P\}$$

As explained below, this captures a notion of indefeasible, absolutely unrevisable knowledge. But we also want a notion of *belief* B, describing "soft" information, which might

[3] A pre-order is any reflexive transitive relation. In Grove's representation theorem the pre-order must also be total and converse-well-founded.

be subject to *revision*. So we want *conditional belief operators* B^P, in order to capture the *revised beliefs* given some new information P. If S is *finite*, let $min_{\leq_i}(P)$ denote the $\leq_i$-minimal P elements $\{s \in P \mid \forall t \in P, s \leq_i t\}$. So $min_{\leq_i}(P)$ denotes the set of states which i considers most plausible given P. Then $min_{\leq_i}(P \cap [s]_i)$ denotes the set of that states which i considers most plausible given both P and i's knowledge at state s. Thus we define player i's *conditional belief* operator as:

$$B_i^Q P := \{s \in S \mid min_{\leq_i}(Q \cap [s]_i) \subseteq P\}.$$

There is a standard way to extend this definition to total preorders on *infinite* sets of states, but we skip here the details, since we are mainly concerned with finite models. $B_i^Q P$ is the event that *agent i believes P conditional on Q*. Conditional belief should be read carefully: $B_i^Q P$ does *not* mean that after learning that Q, i will believe P; rather it means that after learning Q, i will believe that P *was the case before the learning*. This is a subtle but important point: the conditional belief operators do not directly capture the dynamics of belief, but rather as van Benthem [11] puts it, they 'pre-encode' it. We refer to [11, 7] for more discussion. The usual notion of *(non-conditional) belief* can be defined as a special case of this, by putting $B_i P := B_i^S P$. The notions of *common knowledge CkP* and *common belief CbP* are defined in the usual way: first, one introduces *general knowledge* $EkP := \bigcap_i K_i P$ and *general belief* $EbP := \bigcap_i B_i P$, then one can define $CkP := \bigcap_n (Ek)^n P$ and $CbP := \bigcap_n (Eb)^n P$.

It will be useful to associate with the states S some non-epistemic content; for this we use a *valuation function*. Assume given some finite set Φ of symbols, called *basic (or atomic) sentences*, and meant to describe ontic (non-epistemic, non-doxastic) "facts" about the (current state of the) world. A *valuation on* Φ is a function V that associates with each $p \in \Phi$ a set $V(p) \subseteq S$: V specifies at which states p is true. A *plausibility model* for (a given set of atomic sentences) Φ is a plausibility frame equipped with a valuation on Φ.

Interpretation: 'hard' and 'soft' information Information can come in different flavours. An essential distinction, due to van Benthem [11], is between 'hard' and 'soft' information. *Hard information* is absolutely "indefeasible", i.e. *unrevisable*. Once acquired, a piece of 'hard' information forms the basis of the strongest possible kind of knowledge, one which might be called *irrevocable knowledge* and is denoted about by K_i. For instance, the principle of Introspection of Beliefs states that (introspective) agents possess 'hard' information about their own beliefs: they know, in an absolute, irrevocable sense, what they believe and what not. *Soft information*, on the other hand, may in principle be defeated (even if it happens to be correct). An agent usually possesses only soft information about other agents' beliefs or states of mind: she may have beliefs about the others' states of mind, she may even be said to have a kind of 'knowledge' of them, but this 'knowledge' is defeasible: in principle, it could be revised, for instance if the agent were given more information, or if she receives misinformation.

For a more relevant, game-theoretic example, consider extensive games of perfect information: in this context, it is typically assumed (although usually only in an implicit manner) that, at any given moment, both the *structure of the game* and the players' *past moves* are 'hard' information; e.g. once a move is played, all players *know, in an absolute, irrevocable sense*, that it was played. Moreover, past moves (as well as the structure of the game) are common knowledge (in the same absolute sense of knowledge). In contrast, a player's 'knowledge' of other players' rationality, and even a player's 'knowledge' of her own future move at some node that is not yet reached, are not of the same degree of certainty: in principle, they might have to be revised; for instance, the player might make a mistake, and fail to play according to her plan; or the others might in fact play "irrationally", forcing her to revise her 'knowledge' of their rationality. So this kind of defeasible knowledge should better be called 'belief', and is based on players' "soft" information.[4]

In the 'static' setting of plausibility models given above, soft information is captured by the "belief" operator B_i. As already mentioned, this is defeasible, i.e. *revisable*, the revised beliefs after receiving some new information φ being pre-encoded in the conditional operator B_i^{φ}. Hard information is captured by the "knowledge" operator K_i; indeed, this is an absolutely unrevisable form of belief, one which can never be defeated, and whose negation can never be accepted as truthful information. This is witnessed by the following valid identities:

$$K_i P = \bigcap_{Q \subseteq S} B_i^Q P = B_i^{\neg P} \emptyset.$$

Special Case: Conditional Probabilistic Systems If, for each player i, we are given a *conditional probabilistic system* a la Renyi [27] over a common set of states S (or if alternatively we are given a *lexicographic probability system* in the sense of Blume et al), we can define subjective conditional probabilities $\text{Prob}_i(P|Q)$ for events of zero probability. When *S is finite and the system is discrete* (i.e., $\text{Prob}(P|Q)$ is defined for all non-empty events Q), we can use this to define conditional belief operators for arbitrary events, by putting $B_i^Q P := \{s \in S : \text{Prob}_i(P|Q) = 1\}$.

[4] By looking at the above probabilistic interpretation, one can see that the fact that an event or proposition has (subjective) probability 1 corresponds only to the agent having "soft" information (i.e. *believing* the event). "Hard" information corresponds to the proposition being true in *all* the states in the agent's information cell.

It is easy to see that these are special cases of finite plausibility frames, by putting: $s \leq_i t$ iff $\text{Prob}_i(\{s\}|\{s,t\}) \neq 0$. Moreover, the notion of conditional belief defined in terms of the plausibility relation is *the same* as the one defined probabilistically as above.

Dynamics and Information: 'hard' public announcements Dynamic epistemic logic is concerned with the "origins" of hard and soft information: the "epistemic actions" that can appropriately inform an agent. In this paper, we will focus on the simplest case of hard-information-producing actions: *public announcements*. These actions model the *simultaneous joint learning of some 'hard' piece of information* by a group of agents; this type of learning event is perfectly "transparent" to everybody: there is nothing hidden, private or doubtful about it. But dynamic epistemic logic [4] also deals with other, more complex, less transparent and more private, forms of learning and communication.

Given a plausibility model $\mathcal{M} = (S, \leq_i, V)_{i \in N}$ and a "proposition" $P \subseteq S$, the *updated model* $\mathcal{M} \upharpoonright P$ produced by a public announcement of P is given by *conditionalisation*: $(P, \leq_i \upharpoonright P, V \upharpoonright P)$, where $\leq \upharpoonright P$ is the restriction of $\leq$ to P and $(V \upharpoonright P)(p) = V(p) \cap P$. Notice that public announcements can change the knowledge and the beliefs of the players. So far we have, for readability, been writing events without explicitly writing the frame or model in question. However, since we are now talking about model-changing operations it is useful to be more precise; for this we will adopt a modal logical notation.

APAL-CDL: Language and Semantics Our language $APAL - CDL$ is built recursively, in the usual manner, from atomic sentences in Φ, using the Boolean connectives $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$ and $\varphi \Rightarrow \psi$, the epistemic operators $K_i\varphi$, $B_i^{\varphi}\psi$, $Ck\varphi$ and $Cb\varphi$ and the dynamic modalities $[!\varphi]\psi$ and $[!]\varphi$. (The language CDL of conditional doxastic logic consists only of the formulas of $APAL - CDL$ that can be formed *without* using the dynamic modalities.)

For any formula φ of this language, we write $[\![\varphi]\!]_{\mathcal{M}}$ for the *interpretation* of φ, the event denoted by φ, in $\mathcal{M}$. We write $\mathcal{M}^{\varphi}$ for the updated model $\mathcal{M} \upharpoonright [\![\varphi]\!]_{\mathcal{M}}$ after the public announcement of φ. The interpretation map is defined recursively: $[\![p]\!]_{\mathcal{M}} = V(p)$; Boolean operators behave as expected; and the definitions given above of the epistemic operators in terms of events give the interpretation of epistemic formulae. Then the interpretation of the dynamic formulae, which include *public announcement modalities* $[!\varphi]\psi$, goes as follows:

$$[\![[!\varphi]\psi]\!]_{\mathcal{M}} = \{s \in S \mid s \in [\![\varphi]\!]_{\mathcal{M}} \Rightarrow s \in [\![\psi]\!]_{\mathcal{M}^{\varphi}}\}$$

Thus $[!\varphi]\psi$ means that after any true public announcement of φ, ψ holds. The *arbitrary (public) announcement modality* $[!]\varphi$ is to be read: after *every* (public) announcement, φ holds. Intuitively, this means φ is a *"stable" truth*: not only it is true, but it continues to stay true when any new (true) information is (jointly) learned (by all the players). There are some subtleties here: do we require that the new information/announcement be expressible in the language for example? This is the option taken in [3], where the possible announcements are restricted to *epistemic* formulas, and a *complete axiomatisation* is given for this logic. In the context of *finite models* (as the ones considered here), this definition is actually equivalent to allowing *all formulas of our language* $\mathcal{L}$ as announcements. As a result, we can safely use the following apparently circular definition:

$$[\![[!]\varphi]\!]_{\mathcal{M}} = \{s \in S \mid \forall \psi \in \mathcal{L}\ s \in [\![[!\psi]\varphi]\!]_{\mathcal{M}}\}$$

Dynamic epistemic logic captures the "true" dynamics of (higher-level) beliefs after some learning event: in the case of public announcements, the beliefs of an agent i after a joint simultaneous learning of a sentence φ are fully expressed by the operator $[!\varphi]B_i$, obtained by composing the dynamic and doxastic operators. Note that this is *not* the same as the conditional operator B_i^{φ}, but the two are related via the following "Reduction Law", introduced in [11]:

$$[!\varphi]B_i\psi \ \Leftrightarrow\ (\varphi \Rightarrow B_i^{\psi}[!\varphi]\psi).$$

This is the precise sense in which the conditional belief operators are said to "pre-encode" the dynamics of belief.

Special Case: Bayesian Conditioning In the case of a conditional probability structure, the update $\mathcal{M} \upharpoonright P$ by a public announcement $!P$ corresponds to *Bayesian update (conditionalisation)*: the state space is reduced to the event P, and the updated probabilities are given by $\text{Prob}'_i(Q) := \text{Prob}_i(Q|P)$. So a dynamic modality $[!P]Q$ corresponds to the event that, after conditionalising with P, event Q holds. Similarly, the arbitrary announcement modality $[!]P$ is the event that P *stably holds*, i.e. it holds after conditionalising with any true event.

2 Models and languages for games

The notion of *extensive game with perfect information* is defined as usual (cf. [23]): Let N be a set of 'players', and G be a finite tree of 'decision nodes', with terminal nodes (leaves) $\mathcal{O}$ (denoting "possible outcomes"), such that at each non-terminal node $v \in G - \mathcal{O}$, some player $i \in N$ is the decision-maker at v. We write $G_i \subseteq G$ for the set of nodes at which i is the decision-maker. Add to this a payoff function h_i for each player i, mapping all the leaves $o \in \mathcal{O}$ into real numbers, and you have an extensive game. We write 'G' to refer both to the game and to the corresponding set of nodes. We also write $u \to v$ to mean that

v is an immediate successor of u, and $u \rightsquigarrow v$ to mean that there is a path from u to v. A *subgame* of a game G is any game G', having a subset $G' \subseteq G$ as the set of nodes and having the immediate successor relation $\to'$, the set of decision nodes G'_i and the payoff function h'_i (for each player i) being given by restrictions to G' of the corresponding components of the game G (e.g. $G'_i = G_i \cap G'$ etc). For $v \in G$, we write G^v for the subgame of G in which v is the root. A *strategy* σ_i for player i in the game G is defined as usually as a function from G_i to G such that $v \to \sigma_i(v)$ holds for all $v \in G_i$. Similarly, the notions of *strategy profile*, of *the (unique) outcome determined by a strategy profile* and of *subgame-perfect equilibrium* are defined in the standard way. Finally, we define as usually a *backward induction outcome* to be any outcome $o \in \mathcal{O}$ determined by some subgame-perfect equilibrium. We denote by BI_G the set of all backward-induction outcomes of the game G.

Consider as an example the "centipede" game G (cf. [14]) given in Figure 1. This is a two-player game for a (Alice) and b (Bob).

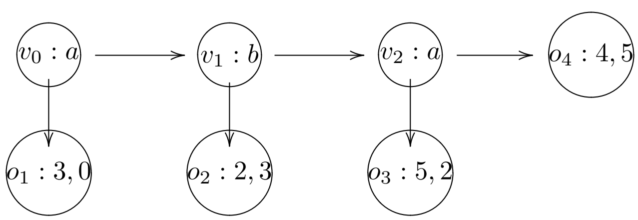

Figure 1. The "centipede" game G

Here, we represent the nodes of the game by circles and the possible moves by arrows. In each circle we write first the name of the node that the circle represents; then, if the node is non-terminal, we write the name of the player who decides the move at that node; while in the terminal nodes (outcomes) o_1, o_2, o_3, o_4, we write the payoffs as pairs (p_a, p_b), with p_a being Alice's payoff, and p_b Bob's. Note that in this game there is one backward induction outcome, o_1, and furthermore that the unique backward induction strategy profile assigns to each v_m the successor o_{m+1}.

Language for Games For any given game G, we define a set of *basic (atomic) sentences* Φ_G from which to build a language. First, we require Φ_G to contain a sentence for each leaf: for every $o \in \mathcal{O}$, there is a basic sentence $\overline{o}$. For simplicity, we often just write o, instead of $\overline{o}$. In addition Φ_G contains sentences to express the *players' preferences over leaves*: for each $i \in N$ and $\{o, o'\} \subseteq \mathcal{O}$, Φ_G has a basic sentence $o \prec_i o'$. Our formal language for games G is simply the language $APAL - CDL$ defined above, where the set of atomic sentences is the set Φ_G. To talk about the non-terminal nodes, we introduce the following abbreviation:

$$\overline{v} := \bigvee_{v \rightsquigarrow o} o\,,$$

for any $v \in G - \mathcal{O}$. As for terminal nodes, we will often denote this sentence by v for simplicity, instead of $\overline{v}$.

Plausibility Models for Games We now turn to defining game models. A *plausibility model for game G* is just a plausibility model $(S, \leq_i, V)_{i \in N}$ for the set Φ_G. We interpret every state $s \in S$ as an *initial state in a possible play* of the game. Intuitively, the sentence $\overline{o}$ is true at a state s if *outcome o will be reached during the play* that starts at s; and the sentence $o \prec_i o'$ says that *player i's payoff at o is strictly smaller than her payoff at o'*.

Observe that nothing in our definition of models for G guarantees that states come with a unique outcome or that the players know the set of outcomes! To ensure this (and other desirable constraints), we later focus on a special class of plausibility models for a game, called "game models".

Examples Figures 2 and 3 represent two different plausibility models $\mathcal{M}_1$ and $\mathcal{M}_2$ for the centipede game G. Here, we use *labelled arrows for the converse plausibility relations* $\geq_a$ (going from less plausible to more plausible states), but for convenience *we skip all the loops*.

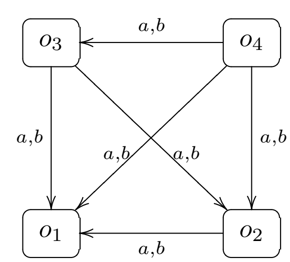

Figure 2. A game model $\mathcal{M}_1$ for the centipede game G

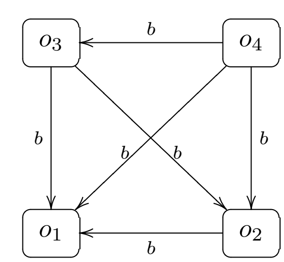

Figure 3. A plausibility model $\mathcal{M}_2$ for G, which is *not* a game model

Note that in the model M_2, Alice (player a) *knows the state of the world*: in each state, she knows both the outcome and Bob's beliefs (and belief revision policy), i.e. the sentence $\bigwedge_{o\in\mathcal{O}}(\mathbf{o} \Rightarrow K_a\mathbf{o})$ holds at all states of M_2. But this is *not true* in model M_1: on the contrary, in M_1 (it is common knowledge that) nobody knows the outcome of the game, and moreover nobody can exclude any outcome. Intuitively, the future is "epistemically open" in M_1, but not in M_2. However, we can also intuitively see that, in both models, (it is common knowledge that) all the players *know the (structure of the) game*: the available outcomes, the structure of the tree, the payoffs etc.

We now want to formalise our intuitions about open future and about having common knowledge of the structure of the game. To do this, we will focus on a special class of models, that we call "game models". Intuitively, each state of a game model comes with a complete play of the game, and hence it should have a uniquely determined outcome, and the set of possible outcomes as well as the players' preferences over them should be common knowledge. However, the players in this (initial) state should not have non-trivial knowledge about the outcome of the play. Indeed, they should have "freedom of choice" during the play, which means they can in principle play any move, so that at the outset of the play they cannot exclude a priori any outcomes.

Game Models The class of *game models for* G, denoted by $\mathfrak{M}_G$, as the class of all plausibility model for G satisfying the following conditions (for all players $i \in N$):

1. $\forall s \in S\, \exists! o \in \mathcal{O} : s \in V(o)$
2. $V(o \prec_i o') = \begin{cases} S & \text{if } h_i(o) < h_i(o') \\ \emptyset & \text{otherwise} \end{cases}$
3. $\forall s \in S\, \forall o \in \mathcal{O} : V(o) \cap [s]_i \neq \emptyset$

The first condition entails that there is common knowledge of the set of possible outcomes, as well as of the fact that each state is associated *a unique actual outcome*. This reflects the fact that the future, for each particular play (state), is determined. The second condition entails that the preferences over outcomes are commonly known. Finally, the third condition says that (it is common knowledge that) the *future is epistemically open*: in the initial state of any play, no player has "knowledge" (in the strong sense of "irrevocable", absolutely unrevisable knowledge) that any outcome is impossible. This is meant to apply even to the states that are incompatible with that player's plan of action.

Open Future We take condition (3) to embody the players' *freedom of choice*, as well as the *possibility of error*: in principle, players might always change their minds or make mistakes, hence any belief excluding some of the outcomes may have to be revised later. Even if we would assume (as usually is assumed) that players *(irrevocably) know their own strategy*, i.e. even if they are not allowed to change their minds, and even if we assume (as postulated by Aumann) that *they have common knowledge of "rationality"* (and so that they can exclude some obviously irrational choices), it still would not follow that they can completely exclude any outcome: mistakes can always happen, or players may always lose their rationality and become temporarily insane; so a rational plan does not necessarily imply a rational play, and hence the future still remains open.

Condition (3) is natural given our interpretation of the "knowledge" operator K as representing *hard information*, that is absolutely certain and irrevocable. If any node is "known" (in this sense) to be unreachable, then that node should simply be deleted from the game tree: this just corresponds to playing a *different game*. So if a player i would irrevocably know that a node is unreachable, then the structure of the game is not "really" common knowledge: i would in fact know that she is playing another game than G. Thus, one can consider the "open future" postulate as a natural strengthening of the "common knowledge of the game" assumption.

A different way to proceed would be to impose the above conditions only locally, at the "real" (initial) state of the play. Let Struct_G be the following sentence, describing the "structure of the game" G:

$$\bigvee_{o\in\mathcal{O}} o \wedge \bigwedge_{o\neq o'\in\mathcal{O}} \neg(o \wedge o') \wedge$$

$$\bigwedge_{\substack{i\in N, o, o'\in\mathcal{O} \\ \text{s.t. } h_i(o)<h_i(o')}} o \prec_i o' \wedge \bigwedge_{\substack{i\in N, o, o'\in\mathcal{O} \\ \text{s.t. } h_i(o)\geq h_i(o')}} \neg o \prec_i o'$$

Similarly, let $\text{F}_G := \bigwedge_{o\in\mathcal{O}, i\in N} \neg K_i \neg o$ be the sentence saying that at the outset of game G the future is epistemically open. Then our proposed "local" requirement is that in the initial state s we have "common knowledge of the structure of the game and of open future", i.e. s satisfies the sentence $Ck(\text{Struct}_G \wedge \text{F}_G)$. Then it is easy to see that this "local" requirement is equivalent to the above global requirement of having a "game model": for every state s in any plausibility model $\mathcal{M}$ for G, s satisfies $Ck(\text{Struct}_G \wedge \text{F}_G)$ iff it is bisimilar[5] to a state in some game model $\mathcal{M}' \in \mathfrak{M}_G$.

Examples Note that the model $\mathcal{M}_1$ from Figure 2 is a *game model*, while $\mathcal{M}_2$ from Figure 3 is *not*: indeed, in $\mathcal{M}_2$ it is common knowledge that Alice always *knows the outcome*, which contradicts the "Open future" assumption.

Encoding Strategies as Conditional Beliefs If a player adopts a particular (pure) *strategy*, our language can encode

[5]Here, "bisimilarity" is the standard notion used in modal logic, applied to plausibility models viewed as Kripke models with atomic sentences in Φ and with relations $\leq_i$. The important point is that our language $APAL-CDL$ cannot distinguish between bisimilar models and states.

this in terms of *the player's conditional beliefs about what she would do at each of her decision nodes*. For instance, we say that Alice "adopts the backward induction strategy" in a given state s of a model for the Centipede Game in Figure 1 iff the sentences $B_a o_1$ and $B_a^{v_2} o_3$ hold at state s. Similarly, we can express the fact that Bob adopts a particular strategy, and by putting these together we can capture *strategy profiles*. A given profile is realized in a model if the correspondent sentence is true at a state of that model.

Note that, in our setting, *nothing forces the players to adopt (pure) strategies*. Strategies are "complete" plans of action prescribing a unique choice (a belief that a particular move will be played) for each decision node of the player. But the players might simply consider all their options as equi-plausible, which essentially means that they do not have a strategy.

Examples In (any state of) model $\mathcal{M}_1$ from Figure 2 it is common knowledge that *both players adopt their backward induction strategies*. In contrast, in the model $\mathcal{M}_3$ from Figure 4, it is common knowledge that *no player has a strategy* (at any node):

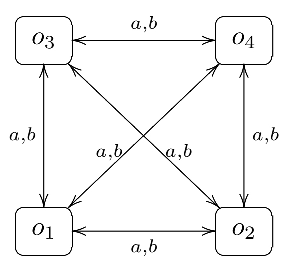

Figure 4. A game model $\mathcal{M}_3$ in which players don't have strategies

So the assumption that *players have (pure) "strategies"* is an extremely strong assumption, which *we will not need*. There is no a priori reason to assume (and there are good empirical reasons to reject) that players play according to fully-determined strategies. Our models are general enough to dispense with this assumption; indeed, our work shows that *this assumption is not needed for proving (common belief) that the backward induction strategy is played*.

Intentions as Beliefs In the above discussion, we identified an agent's *intentions* with *her beliefs about what she is going to do*, and so we represented the decision maker's plan of action as a belief about her (future) action. This identification is philosophically debatable, since agents may be aware of the possibility of mistakes, and so they may doubt that their intentions will be realized. But one can also argue that, in the context of Game Theory, such distinctions will be of very limited significance: indeed, an intention that is not believed to be enforceable is irrelevant for strategic planning (though see [28] for a discussion of intentions in game theory). The players only need to know each others' beliefs about their future actions and about each others' beliefs etc., in order to make their own rational plans; whether or not they are being informed about each others' (completely unenforceable and not believed to be enforceable) "intentions" will not make any difference. So, for the purposes of this paper, we can safely adopt the simplifying assumption that the *agents believe that they will be able to carry out their plans*. Given this assumption, an agent's "intentions" can be captured by her beliefs about her (future) actions.

Representing Players' Evolving Beliefs Recall that we think of every state of a game model $\mathcal{M}_G \in \mathfrak{M}_G$ as an *initial state (of a possible play)* of the game G. As the play goes on, the players' hard and soft information, their knowledge and beliefs, *evolve*. To represent this evolution, we will need to successively change our model, so that e.g. when a node v is reached, we want to obtain a corresponding model of the subgame G^v. That is precisely, in this perfect information setting, what is achieved by *updating the model with public announcements*: indeed, in a game of perfect information, every move, say from a node u to one of its immediate successors u', can be "simulated" by a public announcement $!u'$. In this way, for each subgame G^v of the original model $\mathcal{M}$, we obtain a model $\mathcal{M}^v$, that *correctly describes the players' knowledge and beliefs at the moment when node v is reached during a play*. This is indeed a model of the corresponding subgame G^v:

Proposition 2.1 *If $\mathcal{M} \in \mathfrak{M}_G$ then $\mathcal{M}^v \in \mathfrak{M}_{G^v}$.*

Example Consider a play of the Centipede game G that starts in the initial situation described by the model $\mathcal{M}_1$ in Figure 2, and in which the real state of the world is the one having outcome o_2: so Alice first plays "right", reaching node v_1, and the Bob plays "down", reaching the outcome o_2. The model $\mathcal{M}_1$ from Figure 2 gives us the initial situation, the model $\mathcal{M}_1^{v_1}$ in Figure 5 describes the epistemic situation after the first move, and then the model $\mathcal{M}_1^{o_2}$ in Figure 6 gives the epistemic situation at the end of the play:

In this way, for each given initial state s (of a given play $v_0, v_1, \ldots, o$ of the game, where o is the unique outcome such that $s \in V(o)$), we obtain a *sequence of evolving game models*

$$\mathcal{M} = \mathcal{M}^{v_0}, \mathcal{M}^{v_1}, \ldots, \mathcal{M}^{o},$$

describing *the evolving knowledge and beliefs of the players* during any play. Each model $\mathcal{M}^v$ accurately captures the players' beliefs at the moment when node v is reached. Note also that every such sequence ends with a model $\mathcal{M}^o$ consisting of only one node (a leaf o); this reflects the fact

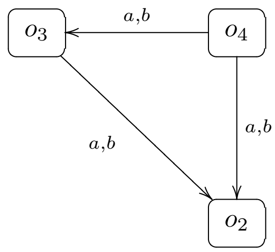

Figure 5. The model $\mathcal{M}_1^{v_1}$

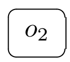

Figure 6. The model $\mathcal{M}_1^{o_2}$

that *at the end of the game, there is no uncertainty left*: the outcome, as well as the whole history of the game, are now common knowledge.

Simulating Moves by Public Announcements Using the dynamic "public announcement" modalities in constructs such as $[!v]B_i$, we can talk, *at the initial state* $s \in \mathcal{M}$ and *without leaving the original model* $\mathcal{M} \in \mathfrak{M}_G$, about all these future, evolving beliefs of the players at nodes v other than the initial node v_0. Indeed, in a game of perfect information, all the moves are *public*. So the epistemic effect of a move to node v is the same as that of a truthful public announcement $!v$ (saying that the node v is reached during the play). In other words, *we can "simulate" moves in games of perfect information by truthful public announcements*.[6]

3 Rationality in Decisions and Games

We now define our fundamental notions of *dynamic rationality* and *rational play*. First we will look at *single-agent (one-step) decision situations*, and then at interactive decision situations, i.e. *games*.

3.1 Single Agent Decision Problems

Given a *one-step decision problem* $\mathcal{P}$ with a set of *outcomes* $\mathcal{O}$, the *decision-maker* i selects one of the outcomes $o \in \mathcal{O}$. The decision-maker may have various hard and soft information about which outcomes can actually be realized and which not. This will determine her knowledge and her beliefs. We assume that *her "hard" knowledge restricts her possible choices*: she *can* only select outcomes that she doesn't *know* to be *impossible*.

[6]We believe that the more general case, of games of imperfect information, can also be handled by using other kinds of epistemic actions proposed in Dynamic Epistemic Logic [4]. But we leave this development for future work.

What this amounts to is the following: for the decision maker i, the "true" set of possible outcomes is $\{o \in \mathcal{O} \mid \neg K_i \neg o\}$, i.e. the set of all the *"epistemically possible" outcomes*. So her selected option must satisfy: $o \in \{o \in \mathcal{O} \mid \neg K_i \neg o\}$. This allows us to capture the "selection" problem using epistemic operators.

To assess whether the decision is "rational" or not, one considers the decision-maker's subjective preferences, modelled as a total preorder $\preccurlyeq_i$ on $\mathcal{O}$. We assume that *agents know their preferences*; indeed, these are interpreted as "doxastic" preferences: beliefs about what's best. Given this interpretation, the CDL postulation of Full Introspection (of beliefs) implies that agents know their preferences.

Rational Choice Rationality, in this case, corresponds to requiring that the selected option is *not worse than* any other (epistemically) possible alternative. In other words, i's solution of the decision problem $\mathcal{P}$ is *rational* if she does not choose any option that is strictly less preferable than an option she doesn't know to be impossible:

$$\mathrm{R}_i^{\mathcal{P}} := \bigwedge_{o,o' \in \mathcal{O}} (o \prec_i o' \wedge \neg K_i \neg o' \Rightarrow \neg o).$$

The main difference between our definition and the standard definition of rational decision-making is the *epistemic limitation of the choice set*. The epistemic operators are used here to delimit *what is currently known about the availability of options*: i's choice should only be compared against options that are not known to be unavailable. This is an *important difference*, and its importance will become clear when we generalise our definition to extensive games.

3.2 Extensive Games

We now aim to extend the above definitions to the case of *multi-agent many-stage decisions*, i.e. "extensive games" (of perfect information). Recall that in an extensive game we are given the players' subjective preferences $\preccurlyeq_i$ only over the leaves. However, at all the intermediate stages of the game, players have to make local choices, not between "final" outcomes, but between "intermediary" outcomes, that is: between other nodes of the game tree.

So, in order to assess players' rationality, we need to *extend* the subjective preference relations to *all the nodes* of the game tree. Fortunately, given the above doxastic interpretation of preferences, there is an obvious (and natural) way to define these extensions. Namely, a player considers a node u to be *strictly less preferable* to a node u' if she *believes* the first to be *strictly dominated* by the second. More precisely, if every outcome that she believes to be achievable given that u is reached is worse than every outcome that she believes to be achievable given that u' is reached:

$$u \prec_i u' := \bigwedge_{o,o' \in \mathcal{O}} (\neg B_i^u \neg o \wedge \neg B_i^{u'} \neg o' \Rightarrow o \prec_i o').$$

By the Full Introspection of beliefs (a postulate of the logic CDL), it follows that we still have that *players know their extended preferences* over all the nodes of the game.

Rationality at a Node Each node $v \in G_i$ can be considered as a (distinct) decision problem, in which the decision-maker is i, the set of outcomes is the set $\{u \in G : v \to u\}$ of all immediate successors of v, and the subjective preference relation is given by the (restriction of the) extended relation $\prec_i$ defined above (to the set $\{u \in G : v \to u\}$). So we can define the rationality of a player i at a node $v \in G_i$ as rationality for the corresponding decision problem, i.e. the player's selection at each decision node consists only of "best answers". Note that, as before, *the player's choice is epistemically limited*: if she has "hard knowledge" excluding some successors (for instance, because those nodes have already been bypassed), then those successors are excluded from the set of possible options. The only difference is that the "knowledge" involved is the one the agent *would have at that decision node*, i.e. it is *conditional on that node being reached*. Formally, we obtain:

$$\mathrm{R}_i^v := \bigwedge_{u,u' \leftarrow v} (u \prec_i u' \wedge \neg K_i^v \neg u' \Rightarrow \neg u)$$

where $K_i^\varphi \psi := K_i(\varphi \Rightarrow \psi)$.

Dynamic Rationality Let R_i be the sentence

$$\mathrm{R}_i = \bigwedge_{v \in G_i} \mathrm{R}_i^v.$$

If R_i is true, we say that player i satisfies *dynamic rationality*. By unfolding the definition, we see it is equivalent to:

$$\mathrm{R}_i = \bigwedge_{v \in G_i} \bigwedge_{u,u' \leftarrow_i v} (u \prec_i u' \wedge \neg K_i^v \neg u' \Rightarrow \neg u).$$

As we'll see, asserting this sentence at a given moment is a way of saying that *the player will play rationally from that moment onwards*, i.e. she will make the best move at any *current or future* decision node.

In the following, "Dynamic Rationality" denotes the sentence

$$R := \bigwedge_i R_i$$

saying that all players are dynamically rational.

Comparison with Substantive Rationality To compare our notion with Aumann's concept of "substantive rationality", we have to first adapt Aumann's definition to a belief-revision context. This has already been done by a number of authors e.g. Battigalli and Siniscalchi [9, 10], resulting in a definition of "rationality at a node" that differs from ours only by *the absence of epistemic qualifications to the set of available options* (i.e. the absence of the term $\neg K_i^v \neg u'$).

The notion of *substantive rationality* is then obtained from this in the same way as dynamic rationality, by quantifying over all nodes, and it is thus equivalent to the following definition:

$$\mathrm{SR}_i = \bigwedge_{v \in G_i} \bigwedge_{u,u' \leftarrow_i v} (u \prec_i u' \Rightarrow \neg u).$$

It is obvious that *substantive rationality implies dynamic rationality*

$$\mathrm{SR}_I \Rightarrow \mathrm{R}_i,$$

but *the converse is in general false*. To better see the difference between SR_i and R_i, recall that a formula being true in a model $\mathcal{M} \in \mathfrak{M}_G$ means that it is true at the first node (the root) of the game tree G. However, we will later have to evaluate the formulas R_i and SR_i at *other* nodes w, i.e. in *other models* of the form M^w (models for *subgames* G^w). Since *the players' knowledge and beliefs evolve during the game*, what is (not) known/believed conditional on v *in model* M^w differs from was (not) known/believed conditional on v *in the original model* (i.e. at the outset of the game). In other words, the meaning of both dynamic rationality R_i and substantive rationality SR_i *will change during a play*. But they *change in different ways*. At the initial node v_0, the two notions are equivalent. But, once a node v has been bypassed, or once the move at v has already been played by a player i, that player is counted as *rational at node* v according to our definition, while according to the usual (non-epistemically qualified) definition the player may have been irrational at v.

In other words, the epistemic limitations we imposed on our concept of dynamic rationality make it into a *future-oriented* concept. At any given moment, the rationality of a player depends only on her current beliefs and knowledge, and so only on the options that she *currently considers possible*: past, or by-passed, options are irrelevant. Dynamic Rationality simply expresses the fact that the player's decision in any future contingencies is rational (given her future options and beliefs). Unlike substantive rationality, our concept has nothing to do with the past or with contingencies that are known to be impossible: a player i may still be "rational" in our sense at a given moment/node v even when v could only have been reached if i has already made some "irrational" move. The (knowledge of some) past mistake(s) may of course affect the others' *beliefs* about this player's rationality; but it doesn't directly affect her rationality, and in particular it doesn't automatically render her irrational.

Solving the BI Paradox As explained above, our concept is very different from (and, arguably, more realistic than) Aumann's and Stalnaker's substantive rationality, but also from other similar concepts in the literature (for example Rabinowicz's [25] "habitual" or "resilient" rationality, etc). The difference becomes more apparent if we consider the assumption that *"rationality" is common belief, in the strongest possible sense,* including *common "strong" belief* (in the sense of Battigalli and Siniscalchi [10]), common persistent belief, or even common "knowledge" in the sense of Aumann. As correctly argued by Stalnaker and Reny, these assumptions, if applied to the usual notions of rationality in the literature, bear no relevance for what the players would do (or believe) at the nodes that are incompatible with these assumptions! The reason is that, if these counterfactual nodes were to be reached, then by that time the belief in "rationality" would have already been *publicly disproved*: we cannot even entertain the possibilities reachable by irrational moves except by suspending our belief (or "knowledge") in rationality. Hence, the above assumptions cannot tell us anything about the players' behaviour or rationality at such counterfactual nodes, and thus they cannot be used to argue for the plausibility of the backward induction solution (even if they logically imply it)! In contrast, our notion of dynamic rationality is *not* automatically disproved when we reach a node excluded by common belief in it: a player *may* still be rational with respect to her current and future options and decisions *even after* making an "irrational" move. Indeed, the player may have been playing irrationally in the past, or may have had a moment of temporary irrationality, or may have made some mistakes in carrying out her rational plan; but she may have recovered now and may play rationally thereafter. Since our notion of rationality is future-oriented, no information about past moves will necessarily and automatically shatter belief in rationality (although of course it *may* still shatter it, or at least weaken it). So it is perfectly consistent (although maybe not always realistic) to assume that players maintain their common belief in dynamic rationality despite all past failures of rationality. In fact, *this is our proposed solution to the BI paradox*: we will show that such a "stable" common belief in dynamic rationality (or more precisely, common *knowledge* of the stability of the players' common belief in rationality) is exactly what is needed to ensure common belief in the backward induction outcome!

Rational Planning A weaker condition requires only that, for each decision node v, the option that the decision-maker *is planning at v to select (at v)* is the best, given the other (epistemically) possible alternatives. By identifying as above the players' plans of actions with their beliefs about their actions, we can thus say that a decision maker is a *rational planner* in the game G if at each decision node she *believes* that she will take "the best decision", even if in the end she may accidentally make a wrong choice:

$$\mathrm{RP}_i := \bigwedge_{v \in G_i} B_i^v \mathrm{R}_i^v .$$

By unfolding the definition, we see it is equivalent to:

$$\mathrm{RP}_i = \bigwedge_{v \in G_i} \bigwedge_{u,u' \leftarrow_i v} (u \prec_i u' \wedge \neg K_i^v \neg u' \Rightarrow B_i^v \neg u).$$

No Mistakes As noted above, $RP_i^{\mathcal{P}}$ only states that the decision maker i has a *rational plan* for current and future contingencies. But mistakes can happen, so if we want to ensure that the decision that is actually taken is rational we need to require the player makes *no mistakes* in carrying out her plan:

$$\text{No-Mistakes}_i := \bigwedge_{v \in G_i} \bigwedge_{u \leftarrow v} (B_i^v \neg u \Rightarrow \neg u)$$

The sentence No-Mistakes$_i$ says that player i's decision are always consistent with her "plan": she never plays a move that, at the moment of playing, she believed won't be played.

As expected, the conjunction of "rational planning" and "no mistakes" entails "rational playing":

$$\mathrm{RP}_i \wedge \text{No-Mistakes}_i \Rightarrow \mathrm{R}_i .$$

4 Backward Induction in Games of Perfect Information

It is easy to see that Aumann's theorem can be strengthened to the following

Proposition 4.1 *In any state of any plausibility model for a game of perfect information, common knowledge of dynamic rationality implies the backward induction outcome.*

Unfortunately, common knowledge of (either dynamic or substantive) rationality *can never hold in a game model*: it is simply incompatible with the "Epistemically-Open Future" condition. By requiring that players have "hard" information about the outcome of the game, Aumann's assumption *does not allow them to reason hypothetically or counterfactually about other possible outcomes*, at least not in a consistent manner.[7] This undermines the intuitive rationale behind the backward induction solution, and it is thus open to Stalnaker's criticism.

So in this section, we are looking for natural conditions that can be satisfied on game models, but that still imply the backward induction outcome (or at least common belief in it). One such condition is *common knowledge of (general) stable belief* in (dynamic) rationality: $Ck[!]EbR$. This is in fact a "strong" form of common belief, being equivalent to $Ck[!]CbR$, i.e. to *common knowledge of stable common belief* in rationality.

[7] Indeed, if o is the backward induction outcome, then the above Proposition entails $K_i o$ for all players i, and thus for every other outcome $o' \neq o$ and every proposition P, we have $B_i^{o'} P$: the players *believe everything* (including inconsistencies) conditional on o'.

Theorem 4.2 *The following holds in any state s of any game model $\mathcal{M} \in \mathfrak{M}_G$:*

$$Ck[!]Eb\mathrm{R} \;\Rightarrow\; Cb(\mathit{BI})\,,$$

where $\mathit{BI} := \bigvee\{o \mid o \in BI_G\}$ is the sentence saying that the current state determines a backward-induction outcome. Equivalently, the following formula is valid over plausibility frames for the game G:

$$Ck(\mathit{Struct}_G \wedge F_G \wedge [!]Cb\mathrm{R}) \;\Rightarrow\; Cb(\mathit{BI}).$$

In English: assuming common knowledge of the game structure and of open future, if it is common knowledge that, no matter what new (truthful) information the players may (jointly) learn during the game (i.e. *no matter what is played*), general belief in rationality will be maintained, then it is common belief that the backward induction outcome will be reached. If we define "stable common belief" in a proposition P as $[!]CbP$, then we can give a more concise English formulation of the above theorem: *common knowledge of the game structure, of open future and of stable common belief in dynamic rationality implies common belief in the backward-induction outcome.*

Although rationality cannot be common knowledge in a game model, *rational planning can be*. When this is the case, we obtain the following

Corollary 4.3 *In a game model, common knowledge of "rational planning" and of stable belief in "no mistakes" implies the backward-induction outcome; i.e. the formula*

$$Ck(\mathrm{RP} \wedge [!]Eb\text{No-Mistakes}) \;\Rightarrow\; Cb(\mathit{BI})$$

is valid on game models.

The above results only give us *common belief* in the backward-induction outcome, but nothing ensures that this belief is correct. If we want to ensure that the backward-induction outcome is actually played, we need to add the requirement that the (stable common) belief in rational play assumed in the premise is correct, i.e. that *players actually play rationally*:

Theorem 4.4 *The following holds in any state s of any game model $\mathcal{M} \in \mathfrak{M}_G$:*

$$\mathrm{R} \wedge Ck[!]Eb\mathrm{R} \;\Rightarrow\; \mathit{BI}$$

No strategies! Observe that *we did not assume that the players have complete (pure) "strategies"* (fully determined plans of action, uniquely specifying one move for at each decision node), but only that they have *partial* plans, i.e. (incomplete) beliefs about what moves should they play: at each decision node they choose a *set* of moves rather than one unique move. So an important side-result of our work is that the assumption that players have (complete, pure) strategies is *not necessary* for proving backward-induction results.

Ensuring Backward-Induction Strategy Profile If, however, we want to postulate that *every player does have a (complete, pure) strategy*, we need to say that, for each node v of her choice, there exists a (unique) immediate successor u that she *believes* will be played if v is reached (i.e. she *plans to play u at v*):

$$\text{Strategies} := \bigwedge_i \bigwedge_{v \in G_i} \bigvee_{u \leftarrow_i v} B_i^v u.$$

In cases where Str is common knowledge as well, we can strengthen the Theorem 4.2 to:

Corollary 4.5 *The following holds in any state s of any game model $\mathcal{M} \in \mathfrak{M}_G$:*

$$Ck(\mathit{Strategies} \wedge [!]Eb\mathrm{R}) \;\Rightarrow\; Cb(\mathit{BI\text{-}Profile})$$

where BI-Profile is the sentence saying that the strategies given by each player's conditional beliefs in the initial state s form a backward-induction profile.

Finally, the following theorem ensures that above results are not vacuous:

Theorem 4.6 *For every extensive game G, there is a game model $\mathcal{M} \in \mathfrak{M}_G$ and a state $s \in \mathcal{M}$ satisfying the sentence*

$$\text{No-Mistakes} \wedge Ck(\mathrm{RP} \wedge \mathit{Strategies} \wedge [!]Eb\text{No-Mistakes}).$$

As a consequence, the sentence $R \wedge Ck[!]Eb\mathrm{R} \wedge Ck\mathit{Strategies}$ is also satisfied.

The *proofs* of these theorems are in **Appendix 1**. Alternative (weaker) conditions ensuring the backward induction outcome are given in **Appendix 2**.

5 Comparison with Other Work

The game-theoretic issues that we deal with in this paper originate in the work of Aumann [2], Stalnaker [30, 31, 32] and Reny [26], and have been investigated by a number of authors [14, 15, 13, 8, 9, 10, 17, 18, 22, 29, 19] etc. Our work obviously owes a great deal to these authors for their illuminating discussions of the topic.

The logic CDL of conditional belief was first introduced and axiomatised by Board [16], in a slightly more complicated form. The version presented here is due to Baltag and Smets [5, 7]. The dynamic extension of CDL obtained by adding the public announcements modalities (coming from the public announcement logic PAL, originally developed by Plaza [24]) has been developed by van Benthem [11] and, independently, by Baltag and Smets [5]. The extension of PAL with arbitrary announcement modalities $[!]\varphi$ is due to Balbiani et al [3]. The belief-revision-friendly version of APAL presented here (obtained by combining APAL with CDL) is an original contribution of our paper.

The work of Battigalli and Siniscalchi [10] is the closest to ours, both through their choice of the basic setting for the "static logic" (also given by conditional belief operators) and through the introduction of a strengthened form of common belief ("common strong belief") as an epistemic basis for a backward-induction theorem. Strong belief, though different from our "stable" belief, is another version of persistent belief: belief that continues to be maintained unless and until it is contradicted by new information. However, their notion of rationality is only "partially dynamic": although taking into account the dynamics of beliefs (using conditional beliefs given node v to assess the rationality of players' choices at v), it does not fully take into account the limitations posed to the set of possible options by the dynamics of "hard knowledge". In common with most other previous notions of rationality, it requires agents to make rational choices at all nodes, including the past ones and the ones that have already been bypassed. As a result, it is enough for a player to make only one "irrational" move to completely shatter the (common) belief (however strong) in rationality; and as a consequence, common strong belief in rationality does not by itself imply backward induction. To obtain their theorem, Battigalli and Siniscalchi have to add another assumption: that the game model is a *complete type structure*, i.e. it contains, in a certain sense, every possible epistemic-doxastic "type" for each player. This means that the players are assumed to have *absolutely no "hard" information, not only about the outcomes or about the other players' strategies, but also about the other players' beliefs*, so that they have to consider as epistemically possible *all* consistent (probabilistic) belief assignments for the other players! This is an extremely extremely strong (and, in our opinion, unrealistic) "completeness" assumption, one that can only be fulfilled in an infinite model. In contrast, the analogue completeness assumption in our approach is the much weaker "Open Future" assumption, postulating that (at the beginning of the game) players have no non-trivial "hard" information about the outcomes (except the information given by the structure of the game): they cannot foretell the future, cannot irrevocably know the players' freely chosen future moves (though they *do* irrevocably know the past, and they *may* irrevocably know the present, including all the beliefs and the plans of action of all the players). Our more realistic postulate is weak enough to be realized on finite models. In particular, it can be realized on models as small as the set of terminal nodes of the game tree (having one state for each terminal node), and in which all the plans of action are common knowledge, so that the only uncertainty concerns possible mistakes in playing (and hence the final outcome).

Samet [29] introduces a notion of *hypothetical knowledge*, in order to develop an epistemic characterisation of backward induction. Hypothetical knowledge looks prima facie similar to conditional belief, except that the interpretation of the hypothetical knowledge formula $K_i^{\varphi}\psi$ is different: "Had φ been the case, i would have known ψ" (op. cit., p. 237). This mixture of counterfactual conditionals and knowledge is specifically introduced in [29] only to discuss backward induction, and it has not occurred before or subsequently in the literature. In contrast, our approach is grounded in the relatively standard and well-understood foundations of Conditional Doxastic Logic, independently studied by logicians and philosophers. While Samet does make what we agree is the important point that some form of counterfactual reasoning is of vital importance to the epistemic situation in extensive games, his model and conditions seem to us more complex, less transparent and less intuitive that ours.

We are aware of only one prior work that uses dynamic epistemic logic (more precisely, the logic of public announcements, but in the context of "classical DEL", i.e. dealing only with knowledge update and not with belief revision) for the analysis of solution concepts in extensive games: van Benthem's work [12]. That work takes Aumann's "static" notion of rationality as given, and accepts Aumann's classical result as valid, and so it does not attempt to deal with the cases in which Aumann's assumptions do not apply, nor to address the criticism and the issues raised by Stalnaker, Reny and others. Instead, van Benthem's contribution focuses on the *sources* of knowledge, on explaining *how* complex epistemic conditions of relevance to Game Theory (such as Aumann's common knowledge of rationality) can *be brought about*, via repeated public announcements of rationality. So van Benthem does *not* use public announcements in order to simulate a play of the game. Public announcements in van Benthem's approach represent *off-line learning*, i.e. *pre-play* or *inter-play* learning, whereas the public announcements in our present approach represent *on-line learning*, i.e. learning that takes place *during the play* of the game. A very interesting open question is to address the same issue answered by van Benthem, but for the case of the dynamic-epistemic condition proposed here, instead of Aumann's condition: find some off-line communication or learning protocol that can achieve common knowledge of stable common belief in rational play.

References

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[2] Robert Aumann. Backward induction and common knowledge of rationality. *Games and Economic Behavior*, (8):6–19, 1995.

[3] Philippe Balbiani, Alexandru Baltag, Hans P. van Ditmarsch, Adreas Herzig, Tomohiro Hoshi, and Tiago de Lima. 'knowable' as 'known after an announcement'. 2008. To appear in Review of Symbolic Logic.

[4] Alexandru Baltag, Lawrence S. Moss, and Slawomir Solecki. The logic of public announcements, common knowledge and private suspicions. Technical Report SEN-R9922, Centrum voor Wiskunde en Informatica, 1999.

[5] Alexandru Baltag and Sonja Smets. Conditional doxastic models: A qualitative approach to dynamic belief revision. *Electronic Notes in Theorerical Computer Science*, 165:5–21, 2006.

[6] Alexandru Baltag and Sonja Smets. The logic of conditional doxastic actions. To appear in Texts in Logic and Games, 2008.

[7] Alexandru Baltag and Sonja Smets. A qualitative theory of dynamic interactive belief revision. To appear in Texts in Logic and Games, 2008.

[8] Pierpaolo Battigalli. On rationalizability in extensive games. *Journal of Economic Theory*, 74(1):40–61, May 1997.

[9] Pierpaolo Battigalli and Marciano Siniscalchi. Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory*, 88(1):188–230, September 1999.

[10] Pierpaolo Battigalli and Marciano Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391, October 2002.

[11] Johan van Benthem. Dynamic logic for belief revision. *ILLC Prepublication*, PP-2006(11), 2006.

[12] Johan van Benthem. Rational dynamics and epistemic logic in games. *International Game Theory Review*, 9(1):13–45, 2007. (Erratum reprint, 9(2), 377–409).

[13] Cristina Bicchieri. Self-refuting theories of strategic interaction: a paradox of common knowledge. *Erkenntnis*, 30:69–85, 1989.

[14] Ken Binmore. Modeling rational players, part I. *Economics and Philosophy*, 3:179–214, 1987.

[15] Ken Binmore. A note on backward induction. *Games and Economic Behavior*, 17(1):135–137, November 1996.

[16] Oliver Board. Dynamic interactive epistemology. *Games and Economic Behavior*, 49:49–80, 2002.

[17] Giacomo Bonanno. The logic of rational play in games of perfect information. *Economics and Philosophy*, 7:37–65, 1991.

[18] Adam Brandenburger. The power of paradox: some recent developments in interactive epistemology. *International Journal of Game Theory*, 35(4):465–492, April 2007.

[19] Thorsten Clausing. Doxastic conditions for backward induction. *Theory and Decision*, 54:315–336, 2003.

[20] Jelle D. Gerbrandy and Willem Groeneveld. Reasoning about information change. *Journal of Logic, Language, and Information*, 6:147–169, 1997.

[21] Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17(2):157–170, 1988.

[22] Joseph Y. Halpern. Substantive rationality and backward induction. 1998.

[23] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[24] Jan A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[25] Wlodek Rabinowicz. Grappling with the centipede: defense of backward induction for bi-terminating games. *Philosophy and Economics*, 14:95–126, 1998.

[26] Philip Reny. Rationality in extensive form games. *Journal of Economic Perspectives*, (6):92–100, 1992.

[27] Alfréd Rényi. On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6:285–335, 1955.

[28] Olivier Roy. *Thinking before acting: intentions, logic, rational choice*. PhD thesis, ILLC, Amsterdam, 2008.

[29] Dov Samet. Hypothetical knowledge and games with perfect information. *Games and Economic Behavior*, (17):230–251, 1996.

[30] Robert C. Stalnaker. On the evaluation of solution concepts. *Theory and Decision*, 37:49–73, 1994.

[31] Robert C. Stalnaker. Knowledge, beliefs and counterfactual reasoning in games. *Economics and Philosophy*, 12:133–163, 1996.

[32] Robert C. Stalnaker. Belief revision in games: forward and backward induction. *Mathematical Social Sciences*, 36:31–56, 1998.

[33] Robert C. Stalnaker. On logics of knowledge and belief. *Philosophical Studies*, 128:169–199, 2006.

Appendix 1: Some Proofs

Definition 5.1 For a finite set $\mathcal{O}$ of "outcomes" and a finite set P of "players", we denote by $\mathfrak{Games}(\mathcal{O}, P)$ the class of all perfect information games having any subset of $\mathcal{O}$ as their set of outcomes and having any subset of P as their set of players. $\triangleleft$

Definition 5.2 A sentence is *valid on a game* G if it is true at every state s of every game model $\mathcal{M} \in \mathfrak{M}_G$.

A sentence is *valid over* $\mathfrak{Games}(\mathcal{O}, P)$ if it is valid on every game $G \in \mathfrak{Games}(\mathcal{O}, P)$. $\triangleleft$

Lemma 5.3 *For every perfect information game G, if we denote the root of G by v_0, the first player of G (playing at v_0) by i and the first move of i (the successor node played at v_0) by v_1, then the sentence*

$$R_i^{v_0} \wedge \bigwedge_{u \leftarrow v_0} B_i^u[!u]BI \wedge [!v_1]BI \Rightarrow BI$$

is valid on G.

Proof. This follows directly from the definition of rationality at a node and the definition of BI. The assumption that $B_i^u[!u]$BI is true at s means that all the states (deemed as "most plausible by i conditional on u") in the set $s_i^u := min_{\leq_i}(\overline{u} \cap [s]_i)$ have only outcomes that are backward induction outcomes in the corresponding subgame: i.e. we have $o(t) \in BI_{G^u}$ for all $t \in s_i^u$. Given that all these outcomes $\{u : u \leftarrow v_0\}$ are consistent with i's knowledge (since we are in a game model), the fact that i is rational at v_0 implies that the successor node v_1 chosen by i must be one that maximises her payoff $h_i(o(s_i^u))$ among all the outcomes in $\bigcup_{u \leftarrow v_0} BI_{G^u}$. But, by the definition, such a node v_1 is exactly the choice prescribed at v_0 by the backward induction strategy! Given this backward-induction choice (v_1) of i at node v_0, and given the fact (ensured by the condition $[!v_1]$BI) that starting from node v_1 everybody *will* play the backward induction choices, we can conclude that the outcome $o(s)$ belongs to the backward induction set of outcomes $BI_{G^v} = BI_G$ for the game G. Hence s satisfies BI. QED

The *Main Lemma* underlying our results is the following:

Lemma 5.4 ***("Main Lemma")*** *Fix a finite set $\mathcal{O}$ of outcomes and a finite set P of players. Let ϕ be any sentence in our language $APAL - CDL$ having the following property: for every game $G \in \mathfrak{Games}(\mathcal{O}, P)$, if we denote the root of G by $v := v_0^G$, the first player of G (playing at v_0) by $i := i_0^G$ and the first move of i (the successor node played at v_0^G) by $v_1 := v_1^G$, then the sentence*

$$\phi \Rightarrow R_i^{v_0} \wedge \bigwedge_{u \leftarrow v_0} B_i^u[!u]\phi \wedge [!v_1]\phi$$

is valid on G.

Under this assumption, we have that the sentence

$$\phi \Rightarrow BI$$

is valid over $\mathfrak{Games}(\mathcal{O}, P)$.

Proof. We need to prove that, for every game $G \in \mathfrak{Games}(\mathcal{O}, P)$, the sentence $\phi \Rightarrow$ BI is valid on G. The proof is *by induction on the length of the game* G.

For games of length 0 (only one outcome, no available moves), the claim is trivial (since the only possible outcome is by definition the backward induction outcome).

Let G be now a game of length $n > 0$, and assume the claim is true for all games of smaller length. Let v_0 be the root of G, i be the first player of G, $\mathcal{M} \in \mathfrak{M}_G$ be a game model for G and s be a state in $\mathcal{M}$ such that $s \models_{\mathcal{M}} \phi$.

Let u be any *arbitrary immediate successor* of v_0 (i.e. any node such that $u \leftarrow v_0$). By the property assumed in the statement of this Lemma, we have that $s \models_{\mathcal{M}} B_i^u[!u]\phi$, and so (if s_i^u is the set defined in the proof of the previous Lemma, then) we have $t \models_{\mathcal{M}} [!u]\phi$ for all $t \in s_i^u$. Hence, we have $t \models_{\mathcal{M}^u} \phi$ for all $t \in s_i^u \cap \overline{u}$. By the induction hypothesis, we must have $t \models_{\mathcal{M}^u}$ BI (since $\mathcal{M}^u$ is a game model for G^u, which has length smaller than G, and so the implication $\phi \Rightarrow$ BI is valid on $\mathcal{M}^u$), for all $t \in s_i^u \cap \overline{u}$. From this we get that $t \models_{\mathcal{M}} [!u]$BI for all $t \in s_i^u$, and hence that $s \models_{\mathcal{M}} B_i^u[!u]$BI.

Let v_1 be now the first move of the game in state s (i.e. the unique immediate successor $v_1 \leftarrow v_0$ such that $s \models_{\mathcal{M}} v_1$). By the property assumed in this Lemma, we have that $s \models_{\mathcal{M}} [!v_1]\phi$. By the same argument as in the last paragraph, the induction hypothesis gives us that $s \models_{\mathcal{M}} [!v_1]$BI. Putting together with the conclusion of the last paragraph and with the fact (following from the theorem's assumption) that $\phi \Rightarrow R_i^{v_0}$ is valid on $\mathcal{M}$, we infer that $s \models_{\mathcal{M}} R_i^{v_0} \wedge \bigwedge_{u \leftarrow v_0} B_i^u[!u]\text{BI} \wedge [!v_1]\text{BI}$. The desired conclusion follows now from Lemma 5.3. QED

Lemma 5.5 *The sentence*

$$\phi \ := \ \text{R} \wedge Ck[!]Eb\text{R}$$

has the property assumed in the statement of Lemma 5.4.

Proof. The claim obviously follows from the following three sub-claims:

1. dynamic rationality is a "stable" property, i.e. the implication $\text{R} \Rightarrow \bigwedge_u [!u]\text{R}$ is valid;
2. the implication $Ck[!]Eb\psi \Rightarrow B_i^u[!u]Ck[!]Eb\psi$ is valid, for all formulas ψ and all nodes $u \in G$;
3. the implication $Ck[!]Eb\psi \Rightarrow [!u]Ck[!]Eb\psi$ is valid, for all formulas ψ and all nodes u.

All these claims are easy exercises in dynamic-epistemic logic. The first follows directly from the definition of dynamic rationality.

The second sub-claim goes as follows: assume that we have $Ck[!]Eb\psi$ at some state of a given model; then we also have $Ck[!u][!]Eb\psi$ for any node u (since $[!]\theta$ implies $[!u][!]\theta$), and so also $K_iCk[!u][!]Eb\psi$ (since common knowledge implies knowledge of common knowledge), from which we get $B_i^uCk[!u][!]Eb\psi$ (since knowledge implies conditional belief under any conditions). This is the same as $B_i^u(u \rightarrow Ck[!u][!]Eb\psi)$, which implies $B_i^u(u \rightarrow Ck^u[!u][!]Eb\psi)$ (since common knowledge implies conditional common knowledge). But this last clause is equivalent to $B_i^u[!u]Ck[!]Eb\psi$ (by the Reduction Law for common knowledge after public announcements).

The third sub-claim goes as follows: assume that we have $Ck[!]Eb\psi$ in some state of a given model; then as before we also have $Ck[!u][!]Eb\psi$, and thus $Ck^u[!u][!]Eb\psi$ (since common knowledge implies conditional common knowledge). From this we get $u \to Ck^u[!u][!]Eb\psi$ (by weakening), which is equivalent to $[!u]Ck[!]Eb\psi$ (by the Reduction Law for common knowledge after public announcements). QED

Theorems 4.4 and 4.2

Proof. Theorem 4.4 follows now from Lemma 5.4 and Lemma 5.5. Theorem 4.2 follows from Theorem 4.4, by applying the operator $Ck[!]Eb$ to both its premiss and its conclusion, and noting that the implication

$$Ck[!]Eb\psi \Rightarrow Ck[!]EbCk[!]Eb\psi$$

is valid. QED

Appendix 2

The epistemic condition $R \wedge Ck[!]EbR$ that was given in this paper (to ensure backward induction) is *not* the weakest possible condition (ensuring this conclusion). Any property ϕ satisfying the assumption of our Main Lemma (Lemma 5.4) would do it. In particular, there exists a *weakest* such condition (the smallest event $E \subseteq S$ such that $E \subseteq R_i^{v_0} \cap \bigcap_{u \leftarrow v_0} B_i^u[!u]E \cap [!v_1]E$), but it is a very complicated and unnatural condition. The one given in the paper seems to be simplest such condition expressible in our language $APAL - CDL$.

However, one can give weaker simple conditions if one is willing to go a bit beyond the language $APAL - CDL$, by adding fixed points for other (definable) epistemic operators.

Let *stable true belief* be a belief that is known to be a stable belief and it is also a stably true belief. Formally, we define:

$$Stb_i\varphi := K_i[!]B_i\varphi \wedge [!]\varphi.$$

Stable true belief is a form of "knowledge", since *it implies truth and belief:*

$$Stb_i\varphi \Rightarrow \varphi \wedge B_i\varphi$$

(and in fact *it implies stable truth*: $Stb_i\varphi \Rightarrow [!]\varphi$). *Knowledge that something is stably true implies stable true belief in it*:

$$K_i[!]\varphi \Rightarrow Stb_i\varphi.$$

Stable true belief is inherently a "positively introspective" attitude, i.e.

$$Stb_i\varphi \Rightarrow Stb_iStb_i\varphi,$$

but it is not positively introspective with respect to ("hard") knowledge:

$$Stb_i\varphi \not\Rightarrow K_iStb_i\varphi.$$

Stable true belief is *not* negatively introspective, neither inherently nor with respect to knowledge.

We can define *common stable true belief* in the same way as common knowledge: first define *general stable true belief*

$$Estb\varphi = \bigwedge_{i \in P} Stb_i\varphi$$

("everybody has stable true belief"), then put

$$Cstb\varphi = \bigwedge_n (Estb)^n\varphi.$$

Note that this definition, although semantically meaningful, is not a definition in our language $APAL - CDL$, since it uses infinite conjunctions. Indeed, we conjecture that common stable true belief is undefinable in the language $APAL - CDL$, since it doesn't seem to be expressible as a combination of common knowledge, common belief and dynamic operators.

Lemma 5.6 *The sentence $Cstb$R satisfies the assumptions of our Main Lemma (Lemma 5.4).*

As an immediate consequence, we have:

Theorem 5.7 The sentence

$$Cstb\mathrm{R} \Rightarrow \mathrm{BI}$$

is valid over game models. In English: (if we assume common knowledge of the structure of the structure of the game and of open future, then) *common stable true belief in (dynamic) rationality implies the backward induction outcome* .