Horacio Arló-Costa
Cristina Bicchieri

Knowing and Supposing in Games of Perfect Information

Abstract. The paper provides a framework for representing belief-contravening hypotheses in games of perfect information. The resulting *t-extended information structures* are used to encode the notion that a player has the disposition to behave rationally at a node. We show that there are models where the condition of all players possessing this disposition at all nodes (under their control) is both a necessary and a sufficient for them to play the backward induction solution in centipede games. To obtain this result, we do not need to assume that rationality is commonly known (as is done in [Aumann (1995)]) or commonly hypothesized by the players (as done in [Samet (1996)]). The proposed model is compared with the account of *hypothetical knowledge* presented by Samet in [Samet (1996)] and with other possible strategies for extending information structures with conditional propositions.[1]

Keywords: Game Theory, Hypothetical Knowledge, Conditionals, Common Knowledge.

1. Introduction: Extending Information Structures

Information structures are a standard representational tool in the theory of games.[2] Structures of this kind appeal to *partitions* of a given space in order to define knowledge. Let us start with a brief review of the partitional account of (unconditional) knowledge. An information structure, for a set of players I, is a list:

$$((\Omega, \Pi_i\ )_{i \in I})$$

where Ω is a set of primitive *states*, and for each player i, Π_i is a partition of Ω.[3] In game theory $\Pi_i(\mathrm{w})$ denotes the unique element of Π_i containing w. We can now introduce the knowledge operators $K_1, \ldots, K_n$, mapping subsets of Ω to subsets of Ω as follows:

[1] This research was funded by the National Science Foundation: "Game-Theoretic Foundation for Multi-Agent Systems", IIS 9734923.

[2] See, for example, chapter 3 of [Rubinstein (1997)] for an introduction to their use in models of perfect and 'bounded' rationality.

[3] We take states as primitives, following the standard practice in modal logic. Points, nevertheless can be given intended interpretations depending on the field of application. In game theory, some possible interpretations of a primitive state may be: a path, pairs of beliefs and strategies, pairs of conditional beliefs and strategies, etc.

Special Issue Formal Epistemology II. *Edited by* **Branden Fitelson**

Studia Logica (2007) 86: 353–373
DOI: 10.1007/s11225-007-9065-6

$$\text{(K) } K_i(\text{E}) = \{\text{w: } \Pi_i(\text{w}) \subseteq \text{E}\}$$

The event K_i(E) is the event that i knows E. Each partition cell P in Π_i can be seen as a possible epistemic state of agent i. Then the event K_i(E) is obtained by taking the union of all the partition cells where E is accepted (i.e. P $\in 2^\Omega$ such that P $\subseteq$ E). Knowledge can then be defined as follows:

$$\text{(k) } K_i(\text{E}) = \cup \{\text{P} \in \Pi_i\text{: P} \subseteq \text{E}\}$$

Dov Samet proposed in [Samet (1996)] to extend this standard framework with an hypothetical operator. An *extended information structure* (EIS) for a set of players I, is a list $((\Omega,\ \Pi_i, T_i\ )_{i \in I})$ where Ω is a set of *states*, and for each player i, Π_i is a partition of Ω. T_i is a *hypothesis transformation* on Π_i. More formally, T_i is a function T_i: $\Pi_i \times (2^\Omega \backslash \emptyset) \rightarrow \Pi_i$, satisfying, for every possible hypothesis H $\in \Omega$:

(T1) T_i(P, H) $\cap$ H $\neq \emptyset$

(T2) If P $\cap$ H $\neq \emptyset$, then T_i(P, H) = P

As we said before, given a player i and a partition Π_i, the partition cells P in Π_i represent the possible epistemic states of i. Suppose that i is in an epistemic state represented by P. Under i's point of view at P, the hypothesis H might either be true, false, or i might be in suspense about its truth-value. Now one can ask: what would i's hypothetical state of mind be, were H to be true? The transformation function defines an hypothetical knowledge operator that gives us i's hypothetical state of mind, were H to be true. Depending on our assumptions about the transformation function, we may end up in very different hypothetical states.

Samet introduced the two postulates listed above, plus the following important constraint (implicit in the definition of the transformation function): Whenever H is not contained in P, the result of hypothesizing H in P leads from P to another cell of Π_i. Samet does not allow for refinements of the original partition.

Knowledge and hypothetical knowledge are then defined as follows: the event K_i(E) is obtained by taking the union of all the partition cells where E is accepted (i.e. Ps such that P $\subseteq$ E). The event that E is known given the hypothesis H, $K_i^H(E)$, can then be constructed by taking all the partition cells P whose transformation with H is covered by E. In other words, we take all the Ps such that E is accepted in T_i(P, H). In Samet's model, the content of a statement of the form: 'Had H been the case, i would have known E' is given by the union of all these partition cells.

$$\text{(K) } K_i^H(E) = \cup \{\text{P} \in \Pi_i\text{: } T_i(\text{P, H}) \subseteq \text{E}\}$$

Samet's model has many virtues as well as some limitations, some of which we considered in [Arlo-Costa and Bicchieri (1998)].[4] Our immediate goal is to gain some insight on the nature and possible use of hypothetical operators by comparing Samet's approach with other standard accounts of transformation functions in the literature. We will then use this analysis in order to propose a new representation of rationality in games of perfect information. Using our characterization of rationality, we prove that much less than common knowledge of rationality is sufficient to determine the backwards induction solution for games of perfect information.

2. Bayesian supposing

An alternative model of hypothetical reasoning is the Bayesian belief revision model. Before the play begins, perhaps the most reasonable representation of an agent's state of mind is given by Ω, the set of primitive states. In other words, the agent has not yet acquired any knowledge about the game; he is in a state of maximal ignorance. He can nevertheless hypothesize that an event occurs and reason about the consequences of such hypothesis. In so doing the agent will change his mind for the sake of the argument, going from some initial state P to an hypothetical scenario T(P, H).[5] The partitional account of knowledge (or conditional knowledge) may not be the best way of modeling this situation. An alternative would be to represent the hypothesis transformation T as a mapping from pairs of propositions to propositions (rather than partition cells to partition cells, given a hypothesis): The first proposition in the pair represents the current epistemic state P (which, in turn, could be hypothetical),[6] the second proposition represents the hypothesis H, and the third proposition represents the resulting hypothetical state, T(P, H).

We may ask of the new transformation function $*_i$: $(2^\Omega \times 2^\Omega) \rightarrow 2^\Omega$, that it satisfies the following typical Bayesian constraints:

(T*1) $*_i(\text{P, H}) \subseteq \text{H}$

(T*2) If $\text{P} \cap \text{H} \neq \emptyset$, then $*_i(\text{P, H}) = \text{P} \cap \text{H}$

(T*3) If $\text{P} \neq \emptyset$, then $*_i(\text{P, H}) \neq \emptyset$.

---

[4]Similar arguments were independently presented in [Halpern (1998)].

[5]If the initial state coincides with the entire space Ω, P is the singleton cell {Ω}.

[6]Consider, for example, the nested transition T(T(Ω, H), H').

(T*1) is essential for a well-behaved notion of supposing. Consider, for example, the case where the agent's state of knowledge is Ω. If in this situation the agent supposes that H is the case (for example, that some node has been reached), it seems that a condition for successful supposing is to move to an hypothetical state where H holds, i.e. where the agent's decision node is indeed reached. Yet the hypothetical process will break down if in the hypothetical scenario the agent continues to be in doubt about whether the node has been reached or not (this will happen, for example, if the hypothetical state continues to be Ω). $*_i$(P, H) can thus be interpreted as answering the question: 'What would the agent know, had he held H true for the sake of the argument?'.

Notice that Samet's theory does not assume (T*1) but the weaker axiom (T1). This is due, in part, to the intended interpretation (in Samet's theory) of T(P, H) as answering the question: 'What would the agent know, had H been true?' — rather than 'What would the agent know, *had H been assumed to be true* for the sake of the argument.'

(T*2) is a Bayesian principle of 'informational economy'. The central idea behind it is 'use conditionalization whenever possible'.[7]

(T*3) is a consistency constraint, assuring that the hypothetical scenarios triggered by consistent propositions are always consistent.

A conditional operator similar to Samet's can be now defined in this Bayesian setting as follows. Given a set of players I, a *dynamic information structure* (DIS) is a list $((\Omega, *_i)_{i\in I})$ where Ω is a set of primitive *states*, and $*_i$ is a *hypothesis transformation* on Ω. $*_i$ is now a function $*_i$: $(2^\Omega \times 2^\Omega) \to 2^\Omega$, satisfying T*1–3 and the following equation (B), which defines a hypothetical knowledge operator:

$$\text{(B) } K_i^H(E) = \cup \{\mathrm{P} \in 2^\Omega \colon *_i(\mathrm{P}, \mathrm{H}) \subseteq \mathrm{E}\}$$

(B) preserves the central definitional features of (K). The main difference with (K) is that (B) allows P to be any arbitrary proposition, whereas (K) restricts the set of possible states of knowledge P to members of Π_i. The postulates T*1–3, in turn, are well grounded in Bayesian theory. It thus seems that the combination of (B) and T*1–3 is exactly what one

[7]The principle does not offer guidance as to how to perform minimal changes prompted by hypotheses contravening current knowledge. In order to do so one needs stronger principles, and in this case it is controversial how to articulate the meaning of 'informational economy'. These issues are beyond the scope of this article. Here we are only concerned with minimal changes prompted by an hypothesis compatible with current knowledge.

needs in order to model hypothetical reasoning in games of perfect information. Unfortunately, we will immediately verify that B and T*1–3 can be jointly held only in trivial models. The result supporting this claim is the game-theoretic counterpart of similar results recently found by philosophical logicians and computer scientists. See chapter 7 of [Gardenfors (1988)] as well as [Lewis (1976)].

Definition 2.1. A hypothetical knowledge operator is *well behaved* if there is no proposition P in Ω, and propositions H, E, and C such that if $C \subseteq \neg E$, $P \subseteq K_i^H(E)$ and $P \subseteq K_i^H(C)$.

The idea behind the notion of a well-behaved operator is simple. We do not want a hypothetical operator to draw contradictory conclusions when an hypothesis is entertained.

Definition 2.2. A *dynamic information structure*, $(\Omega, *_i)_{i\in I}$ is called *non-trivial* if (1) Ω contains three or more states and (2) the hypothetical knowledge operator induced by the structure is *well behaved*.

Claim 2.1. *There are no non-trivial dynamic information structures.*

Proof. Let w_1, w_2, w_3, be three distinct states in the universe Ω of a non-trivial *dynamic information structure*, $(\Omega, *_i)_{i\in I})$. Then we have

$$\{w_1\} \subseteq \{w_1, w_2\} \subseteq K_i^{\{w_2,w_3\}}(\{w_2\})$$

The last inclusion is justified by T*2. In fact, $*_i(\{w_1, w_2\}, \{w_2, w_3\}) = \{w_1, w_2\} \cap \{w_2, w_3\} = \{w_2\}$. But, by the same token, we have:

$$\{w_1\} \subseteq \{w_1, w_3\} \subseteq K_i^{\{w_2,w_3\}}(\{w_3\})$$

Therefore we have a state (w_1) such that $\{w_1\} \subseteq K_i^{\{w_2,w_3\}}(\{w_3\})$ *and* $\{w_1\} \subseteq K_i^{\{w_2,w_3\}}(\{w_2\})$. This contradicts the assumption that the *dynamic information structure* $(\Omega, *_i)_{i\in I}$ is non-trivial. In fact, if $(\Omega, *_i)_{i\in I}$ were non-trivial, the hypothetical knowledge operator K_i should be well-behaved. ■

Results of the previous type are interesting (and perhaps surprising) because they show that two ideas, which are independently coherent and useful, cannot be jointly implemented on pain of triviality. As we said before, (B) *seems* an obvious extension of Samet's ideas, though it requires the use of a transformation function whose interpretation and formal properties diverge from Samet's. On the other hand, the adoption of the postulates T*1–3 is well-grounded in Bayesian ideas. Much of the contemporary work on qualitative theories of belief revision is based on the use of these postulates.

2.1. The Ramsey test

There is an intuitive criterion for acceptance of conditionals that one would like to make compatible with (K) — and also with (B), to the extent to which (B) can be saved by relaxing some of the principles T*1–3. The criterion is that for every permissible state of knowledge P, K_i^H(E) is entailed by P as long as the transformation of P with H entails E. This is known as the *Ramsey test* for conditionals [Ramsey (1990)]. The intuition behind this criterion is that the conditional 'Had H been true (held as true), E would have been the case' should hold in every permissible state of knowledge whose hypothetical revision with H yields E.

In Samet's model, permissible states of knowledge are partition cells, and it turns out that (K) entails, for each partition cell P:

$$\text{(R1) P} \subseteq K_i^H\text{(E) iff } T_i\text{(P,H)} \subseteq \text{E}$$

In Samet's model we also have a corresponding clause for negated conditionals:

$$\text{(R2) P} \subseteq \neg\, K_i^H\text{(E) iff } T_i\text{(P,H)} \not\subseteq \text{E}$$

Is it true that a similar acceptance criterion holds for (B)? Concerning (R1), the question is tantamount to asking whether there is a suitable operator T^i such that it follows from (B) that:

$$\text{(RTP) P} \subseteq K_i^H\text{(E) iff } T^i\text{(P,H)} \subseteq \text{E, for arbitrary propositions P}$$

The answer is yes, but with an important proviso. (RTP) follows from (B), provided that the transformation function T^i obeys a *linearity* property [Collins (1999)].

Claim 2.2. *(RTP) and (B) are equivalent if the transformation operator* T^i *is* linear, *i.e. if for every proposition P, H, and any partition* Π_P *of P and every cell* P_j *of* Π_P*:* $T^i(P, H) = \cup\{T^i(P_j, H)\}$.

The problem is that linearity is incompatible with postulates T*1–3. In other words, when the hypotheses and permissible states of knowledge are not restricted in any special manner,[8] there is no linear transformation function obeying T*1–3 (we leave the verification of this fact to the reader as a simple exercise).

[8]For example, a restriction might require that the permissible states of knowledge are partition cells, which is basically Samet's solution.

(RTP) is well known in the literature. It was proposed by Peter Gardenfors in [Gardenfors (1988)] as a test for acceptance of *ontic* conditionals (the only difference is that Gardenfors uses sentences rather than propositions in order to formulate the test). Therefore the previous claim shows that when the transformation function is linear (i.e. it is an imaging operator) (B) is equivalent to a well-known condition for acceptance of conditionals used in the literature. This provides additional intuitive support for the condition (B) by showing that it is equivalent well-known acceptance conditions for ontic (or subjunctive) conditionals (for linear revisions).

Our first claim, therefore, shows that there is a tension between well-supported Bayesian constraints on transformation functions ($T*2$) and the condition (B) and that this tension is explained in terms of the fact that (B) and (RTP) are equivalent for linear transformations (it is well known that there are no non-trivial theories jointly stating (RTP) and ($T*2$)).

2.2. Circumventing the problem

There are (at least) three possible ways out of the problem we just diagnosed. One option is to keep (RTP) and give up some of the postulates T*1–3. This approach was first suggested by David Lewis [Lewis (1976)]. Lewis faced a probabilistic analogue of our triviality result, and opted for the radical solution of abandoning conditionalization as a rule for supposing. His proposal is to use a new linear rule for updating, called *imaging*. In this setting both T*2 and T*3 are thus weakened:

(T*i2) If P $\subseteq$ H, then P $\subseteq T^i$(P, H)

(T*i3) If P $\neq \emptyset$ and H is consistent, then T^i(P, H) $\neq \emptyset$.

Many of the contemporary versions of *causal decision theory* are based on the substitution of imaging to conditioning in the calculation of expected utility.[9]

Another possible solution is to keep T*1–3 and to give up (RTP). This can be accomplished by developing a syntactic model where hypothetical operators do not express propositions. See 'General Propositions and Causality' in [Ramsey (1990)] for a historical reference, and the section of [Cross and Nute (1998)] devoted to epistemic conditionals for a review of contemporary work in the field. Chapter 7 of [Gardenfors (1988)],

[9]Some authors have recently suggested that the defining ingredient of the notion of supposition used in causal decision theory is exactly *linearity* ([Gardenfors (1988)], [Joyce (199)], [Collins (1999)]).

[Levi (1988)], [Arlo-Costa (1999)], and [Hansson (1992)] offer an overview of the syntactical model of conditionals in games along the lines sketched in [Arlo-Costa and Bicchieri (1998)].

Our solution in this paper (compatible with accepting conditional propositions) is to keep (RTP) together with a restricted version of T*1–3. According to this restriction, as we show in the next section, the transformation function ranges only over primitive states in Ω.

3. Extended information structures revisited

A *t-extended information structure* (t-EIS) for a set of players I, is a list $((\Omega, t_i\ )_{i \in I})$ where Ω is a set of *states*, and for each player i, t_i is a *hypothesis transformation* on Ω. t_i is a function t_i: $\Omega \times (2^{\Omega} \backslash \emptyset) \rightarrow 2^{\Omega}$ satisfying, for all w in 2^{Ω},

(t*1) $t_i(w, \mathrm{H}) \subseteq \mathrm{H}$

(t*2) If $\{w\} \cap \mathrm{H} \neq \emptyset$, then $t_i(w, \mathrm{H}) = \{w\} \cap \mathrm{H}$

(t*3) If $\mathrm{H} \neq \emptyset$, then $T_i(w, \mathrm{H}) \neq \emptyset$.

Intuitively $t_i(w, \mathrm{H})$ can be seen as the set of states of the world considered possible by an agent who supposes that H is true, if her current epistemic state is w. This interpretation should be carefully differentiated from a reading of $t_i(w, \mathrm{H})$ as '$t_i(w, \mathrm{H})$ is the set of states of the world considered possible by an agent who supposes that H is true, if the actual state of the world is w. This ontological reading of $t_i(w, \mathrm{H})$ is quite foreign to the central accounts of belief revision in the existent literature. Neither Gardenfors, Levi or Spohn, just to mention three important authors in the field considers this ontologizing version feasible, and we will here follow standard usage in dismissing such an interpretation of the transformation function.

We can now define a hypothetical knowledge operator as follows:

$$\text{(k) } K_i^H(E) = \{w \in \Omega\text{: } t_i(w, \mathrm{H}) \subseteq \mathrm{E}\}$$

From a technical point of view the transformation function just introduced is a *selection function* of the type used in standard possible worlds semantics for conditionals, rather than a more general suppositional operator capable of modifying not only states of the underlying space, but also sets of states. See [Stalnaker (1968)], [Stalnaker and Thomason (1970)], [Lewis (1973)], [Bicchieri (1988)], [Bicchieri (1989)], and [Bicchieri (1994)] as well as [Cross and Nute (1998)] for suitable references.

Notice that the following postulates can be deduced from the previous ones:

(t1s) $t_i(w, \mathrm{H}) \cap \mathrm{H} \neq \emptyset$

(t2s) If $\{w\} \cap \mathrm{H} \neq \emptyset$, then $t_i(w, \mathrm{H}) = \{w\}$

(t*i2) If $\{w\} \subseteq \mathrm{H}$, then $\{w\} \subseteq t_i(w, \mathrm{H})$

It should be noted that a t-EIS does not appeal to partitions at all. In what follows, we will not use (unconditional) knowledge operators, and we will not need partitions, although nothing precludes adding them to the t-EIS models for some purposes. In other words, our conclusions do not depend upon the type of partitions eventually used to model a game. We only need a set of states and hypothetical transformations of those states.

After this remark we can return to our main line of argument. Notice that $t_i(w, \mathrm{H})$ might select any (not necessarily proper) subset of the *consequent proposition* H. Any such selection is compatible with our axioms. For example, in a centipede game $\{w\}$ might encode the path recommending to go 'down' at the root. Here it is useful to keep in mind that the states in Ω should be understood as *paths.*

$t_i(w, \mathrm{H})$ encodes a hypothesis, made by the player who moves at the root, about what would happen if a successor node H were to be reached. Our axioms are neutral with regard to the content of such a supposition. Some models might allow for $t_i(w, \mathrm{H}) = \mathrm{H}$, which would represent a state of maximal ignorance of the player engaging in the hypothetical exercise. In other words, $t_i(w, \mathrm{H}) = \mathrm{H}$ would include all the paths that pass through node H. Other models might impose stronger requirements. For example, it could be possible to uniquely select a subset of H as $t_i(w, \mathrm{H})$. One such subset may recommend playing the backwards induction strategy once H is reached. Axiomatic constraints in our models do impose, nevertheless, some limitations. For example, H should be known in the hypothetical state $t_i(w, \mathrm{H})$. This, as we explained before, is standard practice in Bayesian models and departs from Samet's model, whose goal is to capture the content of what is known given that a fact *is* true, rather than *supposed* to be true.

To see how t_i-functions work, consider the following three-node centipede. The root is node 1, and then we have successor nodes 2 and 3. Player I plays at the root and at node 3, and player II plays at node 2. w_1, w_2, w_3, w_4 are four possible paths, or complete histories of the game. Each path also represents an epistemic state. For example, w_3 is the epistemic state in which player I knows he plays across at node 1, and player II plays across at node II, and finally I plays down at node 3 (with payoffs (3, 1)). Path w1 consists on player I playing down at root (with payoffs (1, 0)). In path w2, player I plays across and player II down (with payoffs (0, 2)). Finally in path w4 both player I and player II play across. Player I plays across twice and the path ends with payoffs (2, 3). Even if a player is in a specific state of knowledge, he may entertain a hypothesis about what would happen were he to play differently. For example, hypotheses of the following sort would be natural candidates for evaluation by player I: 'I would play down at node 3 if II were to play across at node 2'. Such a conditional can be entertained by player I, even if he plays down at the root and knows it. The relevant transformation function is $t_I(w_1, \mathrm{H})$, where $\mathrm{H} = \{w_3, w_4\}$. If $t_I(w_1, \mathrm{H}) = \{w_3\}$, player I accepts the conditional in question from his epistemic state w_1.

There are, nevertheless, some limitations in the use of transformation functions to track the dynamics of epistemic states. For example, one might want to represent the act of sequentially entertaining two hypotheses. The first establishes that node 2 has been reached ($\mathrm{F} = \{w_2, w_3, w_4\}$). The second evaluates H (saying that player II plays across at node 2), *from F's point of view*. The corresponding nested usage of t_I is:

$$t_I(t_I(w_1, \mathrm{F}), \mathrm{H})$$

Note that this formula is meaningful only if $t_I(w_1, \mathrm{F})$ is a singleton. Notice, nevertheless, that this restriction, which is rather common in equilibrium analysis, is only a restriction about feasible *suppositional* states. It should be stressed that we are not restricting the set of possible epistemic states of players (which could be in epistemic states that are far from omniscient), we are only restricting the set of feasible suppositions made by agents.

t_i is, as we just remarked, a function t_i: $\Omega \times (2^{\Omega}\backslash\emptyset) \rightarrow 2^{\Omega}$. Of course, one might extend the representational power of t_i-functions by defining $t_I(\mathrm{F}, \mathrm{H})$ as the union of $t_I(w_i, \mathrm{H})$, for i ranging in $\{w_2, w_3, w_4\}$. This can be done, at the price of re-introducing a version of the condition we called *linearity*, and this condition will certainly bring us far away from the Bayesian path we have followed so far (notice that such a function will not obey the crucial Bayesian postulate (t*2) presented above). Alternatively one can assume that transformation functions *cumulate* in the sense that $t_I(t_I(w_i, \mathrm{F}), \mathrm{H}) = t_I(w_i, \mathrm{F} \cap \mathrm{H})$, for arbitrary propositions F and H. Some recent dynamic extensions of Harsanyi's theory of types make cumulative assumptions (see, for example [Battigalli and Bonnano (1997)] for a recent analysis of issues related to this point). Cumulativity seems also required in order to guarantee (sufficiently rich) maps between primitive conditional probability (the so-called Renyi-Popper functions used in dynamic extensions of type theory) and infinitesimal probability (see [Arlo-Costa and Thomason (2001)]) (see also [Arlo-Costa (2001)] for a review of the use of cumulative assumptions in probabilistic models of conditionals of the type offered in [Adams (1975)]). Yet cumulativity is not universally assumed as a constraint on supposition (see, [Arlo-Costa and Thomason (2001)], sections 5 to 8) and assuming it is not compatible with an unrestricted version of postulate t*3. We will thus not appeal to this assumption here, even when its use for representing supposition in games of perfect information might be justifiable (at least for models with finite or at most countable universes).

See [Arlo-Costa and Thomason (2001)] and [Arlo-Costa (2001)] for a probabilistic justification of a non-consistency preserving and cumulative t-function. It should be said in passing that less than the full force of cumulativity can perhaps be used in models of games of perfect information. For example, one might require the mutual compatibility of H and F, as well as the fact that both propositions are epistemically possible for player I (see [Arlo-Costa and Thomason (2001)] for a precise definition of epistemic possibility). This weakened version of cumulativity is consistency preserving.

The introduction of conditional propositions in information structures is not easy to reconcile with the idea of providing Bayesian foundations for the theory of games. It is possible to build extended information structures containing such conditional propositions, as long as we accept some constraints on expressive power. The model we present here introduces conditional propositions and sets various constraints of the type exemplified above. Our characterization of rationality does not require either iterated conditionals or the evaluation of hypotheses from the point of view of non-singleton events. On the other hand, even when our transformation functions range over the domain of states Ω, rather than over 2^{Ω}, the suppositional outputs can be arbitrary events. It should be said here in passing that our model answers to criticisms of Samet's model advanced in [Arlo-Costa and Bicchieri (1998)] and [Halpern (1998)]. Moreover, our selection functions have a different interpretation than Samet's ones.

4. Models for games of perfect information

A finite *game* G of perfect information consists of a finite set of players I and a finite tree with a set of nonterminal nodes V, a set of terminal nodes Z, and a root r (we follow Samet's notation here). For each player $i \in$ I,

$V_i \subseteq V$ is the set of i's decision nodes. For two nodes u, v, we write $v \preceq u$ or $u \succeq v$ when u is a node in the subtree the root of which is v. For $v \in V$, A(v) denotes the set {a | (v, a) is an arc of the tree}. The members of A(v), when $v \in V_i$, are called i's actions at v.

A *strategy* for player i is a function s_i: $V_i \rightarrow V \cup Z$, such that for each $v \in V_i$, $s_i(v) \in A(v)$. A strategy profile s is a combination of strategies $s = (s_i)_{i \in I}$. *s(v)* denotes the terminal node that is reached by strategy profile s from node v. Thus, if v itself is a terminal node, then *s(v)* $= v$.

A *model* for a game of perfect information G with set of players I is a pair (E, ζ), where E is an extended information structure and ζ is a map ζ: $\Omega \rightarrow$ Z, onto the set Z of terminal nodes, such that for every player i, node v $\in V_i$, and action a $\in$ A(v), [a] $\subseteq K_i$([v] $\rightarrow$ [a]) where [v] is the event that node v is reached and [a] the event that action a is performed. The player's *payoff function* is a real valued function h_i: Z $\rightarrow$ R. Let the payoff maps η_i on Ω be defined by $\eta_i(\mathrm{w}) = h_i(\zeta(\mathrm{w}))$. In addition the event $\{\mathrm{w}\colon \eta_i(\mathrm{w}) < \mathrm{x}\}$ is denoted by $[\eta_i < \mathrm{x}]$.

The payoffs are assumed to be non-degenerate. Therefore there is a unique strategy $\beta = (\beta_i)_{i \in I}$ satisfying, for each i and $v \in V_i$, the condition $\beta_i(v) = argmax_{a \in A(v)}\ h_i(\beta(a))$. β is called the *backward induction strategy*, and $\beta(r)$ is called the *backward induction path*. According to the previous notation, $\beta(r)$ denotes the terminal node that is reached by the backward induction strategy from the root.

Now we can define the key notion of this section. The event that player i is rational at node v $\in V_i$ is defined as follows:

$$\mathrm{R(v)} = [\mathrm{v}] \cap \cap_x \cap_{a \in A(v)} \neg\, (([\mathrm{v}] \rightarrow [\eta_i < \mathrm{x}]) \cap K_i^{[a]}(\ [\mathrm{a}] \rightarrow [\eta_i \geq \mathrm{x}]))$$

The definition is a slight variation of the characterization of rationality at a node offered by Samet in [Samet (1996)]. Unlike Samet's, our characterization is completely independent of the use of partitions, and therefore we eliminate an extra (unconditional) knowledge operator also used by Samet. It is important to emphasize that the notion just defined is a *behavioral concept*, i.e. rationality at a node v is a concept that *only* applies to states where v is indeed reached. The intuition is that player i is rational at node v if that node is in fact reached and if there is no number x such that i's action at v nets him a payoff less than x, whereas he hypothesizes that another action he might choose at v would net him a payoff of at least x.

One of the main goals of [Samet (1996)] is to find sufficient rationality conditions to guarantee backward induction play. Several obvious conditions do not suffice. For example, Samet shows in [Samet (1996)] that neither common knowledge of rationality at a node, nor the event that players hypothesize that they are rational at each decision node are enough. Samet proposed the notion of *common hypothesis of node rationality*, and proved that it entails the backward induction path. For each pair of nonterminal nodes u and v such that $v \succeq u$, and event E, the event that there is a *common hypothesis of E from u to v*, $H(u, v, E)$, is defined inductively as

$$H(u, u, E) = E$$

If $u \in V_i$, $v \succeq u$, and $H(a, v, E)$ is defined for the (unique) node a in $\mathrm{A}(u)$ on the path from u to v then

$$H(u, v, E) = K^{[a]}([\mathrm{a}] \rightarrow H(a, v, E))$$

The event that there is a *common hypothesis of node rationality* is

$$\bigcap_{v \in V} H(r, v, R(v))$$

5. Rationality at a node and rationality as a disposition

The event that player i is rational can be constructed in terms of a behavioral notion of rationality by identifying it with the event that i hypothesizes that he is rational at all his nodes: $\cap_{v \in V_i} K_i^{[v]}([\mathrm{v}] \rightarrow \mathrm{R(v)})$.

We propose an alternative definition of players' rationality that, when it is distributed knowledge among the players, guarantees the backward induction play. This notion of rationality (*Rat*) is based upon the notion of rationality at a node. Let us say that a player has a *disposition* to be rational if, for every node v at which he might choose, it is not possible that he knowingly chooses an irrational action at v.

$$\mathrm{Rat}(i) = \cap_{v \in V_i} \cap_{a \in A(v)} \neg \, (K_i^{[v]}(\,[\mathrm{v}] \rightarrow [\mathrm{a}]) \cap K_i^{[a]}(\,[\mathrm{a}] \rightarrow \neg \, \mathrm{R(v)})$$

We shall now focus on the class of games known as *centipede games* - defined over binary trees. A game of perfect information is called a centipede game if and only if there is a vertex v whose actions at v lead to terminal nodes and such that for every $u \preceq v$, and actions a, $d \in \mathrm{A}(u)$, one of the actions, say d, leads to a terminal node and $[a] = [s]$, where s is the vertex immediately succeeding u.

THEOREM 5.1. *For each centipede game there exists a model in which:*

$$Rat = [\beta(\mathrm{root})]$$

PROOF. We will split the proof into two main lemmas. First we will show the existence of models such that for each centipede game [β(root)] entails *Rat*. This proof also shows in all generality that for each game of perfect information there exist a model such that *Rat* is non-empty. In the second part of the proof we will show that *Rat* entails [β(root)].

We remind the reader that the payoffs are assumed to be non-degenerate. Therefore there is a unique $\beta = (\beta_i)_{i \in I}$ satisfying, for each i and $v \in V_i$, the condition $\beta_i(v) = argmax_{a \in A(v)}\ h_i(\beta(a))$. It is useful to remember here that if a is a terminal node, $h_i(\beta(a)) = h_i(a)$. β is called the *backward induction strategy*, and $\beta(root)$ is called the *backward induction path*. According to the previous notation, $\beta(root)$ denotes the terminal node that is reached by the backward induction strategy from the root.

In order to show the existence of a model capable of solving centipede games, we will choose a particular transformation function. In this proof we will use the transformation function that selects backward induction solutions in sub-trees.

Define t_i for each path w in Ω and non-empty hypothesis H $\subseteq 2^{\Omega}$, as follows:

If $w \notin$ H, then set $t_i(\{\text{w}\}, \text{H}) = \{[\beta(t)]\}$, for some t which is minimal with respect to $\succeq$ in H. Otherwise $t_i(\{\text{w}\}, \text{H}) = \{\text{w}\}$. It is easy to check that t_i is a transformation function obeying the corresponding axioms. Moreover this definition entails that for each $u \in$ V:

$$\text{(I)}\ [\beta(u)] \subseteq R(u)$$

Samet proves (I) in his paper, but his definition of rationality at a node differs from ours. So we need to re-check this fact here. Let a and d be the actions at node u and i the player playing at u.

Assume first, that $[\beta(u)] \subseteq [a]$. In this case, in virtue of the fact that for each agent i and $v \in V_i$, $\beta_i(v) = argmax_{a \in A(v)}\ h_i(\beta(a))$, we claim that $h_i(\beta(a)) > h_i(\beta(d))$. The proof branches into two sub-cases. First consider the case $x \leq h_i(\beta(u))$. In this case we have:

$$\text{(II)}\ [\beta(u)] \notin (\overline{[u]} \cup [\eta_i < \text{x}])$$

which guarantees that, for each payoff $x \leq h_i(\beta(u))$, and each action a at u: $[\beta(u)] \notin (([\text{u}] \rightarrow [\eta_i < \text{x}]) \cap K_i^{[a]}(\ [\text{a}] \rightarrow [\eta_i \geq \text{x}]))$.

The second sub-case contemplates values of $x > h_i(\beta(u))$. In this case we have:

$$\text{(a) } t_i([\beta(u)], [d]) = [d] \notin ([\mathrm{d}] \to [\eta_i \geq \mathrm{x}])$$

$$\text{(b) } t_i([\beta(u)], [a]) = [\beta(u)] = [\beta(a)] \notin ([\mathrm{a}] \to [\eta_i \geq \mathrm{x}])$$

In order to see that (a) holds it is enough to notice that, given the assumptions, $x > h_i(\beta(u)) = h_i(\beta(a)) > h_i(\beta(d))$. Similar considerations suffice to establish (b). This, in turn, is sufficient to establish the case $[\beta(u)] \subseteq [a]$.

We can now check the case $[\beta(u)] \subseteq [d]$. In this case we have $h_i(\beta(d)) > h_i(\beta(a))$. The proof of this sub-case branches again. When $x \leq h_i(\beta(u))$ condition (II) holds again. On the other hand, when $x > h_i(\beta(u))$, we also have

$$t_i([\beta(u)], [d]) = [d] \notin ([\mathrm{d}] \to [\eta_i \geq \mathrm{x}])$$

and since $x > h_i(\beta(d)) = h_i(\beta(u)) > h_i(\beta(a))$, we also have:

$$t_i([\beta(u)], [a]) = [\beta(u)] = [\beta(a)] \notin ([\mathrm{a}] \to [\eta_i \geq \mathrm{x}])$$

We will show now for an arbitrary node u that:

$$\text{(A) } [\beta(root)] \subseteq Rat(u) = \cap_{a \in A(u)} \neg\, (K_i^{[u]}([\mathrm{u}] \to [\mathrm{a}]) \cap K_i^{[a]}([\mathrm{a}] \to \overline{R(u)})$$

First case: $[\beta(root)] \notin [u]$

In this case $t_i([\beta(root)], [u]) = [\beta(u)]$, for each player i. Call a and d the two actions at u. Assume w.l.o.g. that $[\beta(u)] \in [a]$. In order to establish (A) it is enough to notice that the following two facts hold. First $\beta(root) \notin (K_i^{[u]}([\mathrm{u}] \to [\mathrm{d}]) \cap K_i^{[d]}([\mathrm{d}] \to \overline{R(u)})$, given that:

$$t_i([\beta(root)], [u]) = [\beta(u)] \notin ([\mathrm{u}] \to [\mathrm{d}])$$

On the other hand, since $[\beta(u)] \subseteq R(u)$, $\beta(root) \notin (K_i^{[u]}([\mathrm{u}] \to [\mathrm{a}]) \cap K_i^{[a]}([\mathrm{a}] \to \overline{R(u)})$, in virtue of:

$$t_i([\beta(root)], [a]) = [\beta(u)] = [\beta(a)] \notin ([\mathrm{a}] \to \overline{R(u)}).$$

Therefore $[\beta(root)] \subseteq Rat(u)$ as desired. Now we have to consider the second case:

Second case: $[\beta(root)] \in [u]$

Assume w.l.o.g. that $[\beta(root)] \in [a]$. Then $t_i([\beta(root)], [u]) = t_i([\beta(root)], [a]) = [\beta(root)]$. And we also have $[\beta(root)] = [\beta(u)]$, as well as $[\beta(u)] \in R(u)$. This is enough to guarantee that:

(B) $[\beta(root)] \subseteq Rat(u) = \cap_{a \in A(u)} \neg\, (K_i^{[u]}([\mathrm{u}] \rightarrow [\mathrm{a}]) \cap K_i^{[a]}(\, [\mathrm{a}] \rightarrow \overline{R(u)})$

In fact, as before,

$$t_i([\beta(root)], [u]) = [\beta(root)] \notin ([\mathrm{u}] \rightarrow [\mathrm{d}])$$

and

$$[\beta(root)] \notin ([\mathrm{a}] \rightarrow \overline{R(u)})$$

We will conclude the proof by establishing that *Rat* entails $[\beta(\mathrm{root})]$. Throughout the proof we will call v the node such that there is an action at v leading to a terminal node, say d, such that $[d] = [\beta(\mathrm{root})]$. Assume by contradiction that there is a path $[p] \in Rat$ and $[p] \neq [\beta(\mathrm{root})]$. This proof branches into two main cases. Either there is a node $t \preceq v$, such that there is an action $d_t \in \mathrm{A}(t)$, and $[p] = [d_t]$; or $v \prec t$ (this being the second case). Since we are working with binary trees, we will call a_t the 'across' action at t.

It is enough to establish that $[p] \notin \mathrm{Rat}(t)$. To show this, in turn, it is enough to show that, if player i is playing at node t:

(1) $t_i([p], [t]) = [p] \subseteq \overline{[t]} \cup [d_t]$.

(2) $t_i([p], [d_t]) = [p] = [d_t] \subseteq \overline{[d_t]} \cup \overline{R(t)}$

(1) is obviously true. Establishing (2) requires showing that $[d_t] \notin R(t)$. We know that the backwards induction strategy β is such that $\beta_i(a_t)) = h_i(\zeta([\beta(\mathrm{root})])$. Moreover, $h_i(\zeta([p]) < \beta_i(a_t)) = h_i(\zeta([\beta(\mathrm{root})])$.

In order to show that that $[p] = [d_t] \notin R(t)$ it is sufficient to establish that $[p] \in (\overline{[t]} \cup [\eta_i < \mathrm{x}]) \cap K_i^{[a_t]}([a_t] \rightarrow [\eta_i \geq \mathrm{x}])$. Pick $x = h_i(\zeta([\beta(\mathrm{root})])$. This guarantees that $t_i([p], [a_t]) = [\beta(\mathrm{root})]) \subseteq (\overline{[a_t]} \cup [\eta_i \geq \mathrm{x}])$, given that we have selected $x = h_i(\zeta([\beta(\mathrm{root})])$ - which guarantees $[\beta(\mathrm{root})] \in [\eta_i \geq \mathrm{x}]$. On the other hand $[p] \in [\eta_i < \mathrm{x}]$, guaranteeing that $[p] \notin R(t)$.

Now we have to consider the second case. I.e. we assume by contradiction that there is a node t, which is a successor of v ($v \prec t$), such that there is an action $d_t \in \mathrm{A}(t)$, and $[p] = [d_t] \in Rat$.

There are two main sub-cases to consider. First assume that $[\beta_i(t)] \neq [d_t]$. In this sub-case $[p] \in (K_i^{[t]}([\mathrm{t}] \rightarrow [d_t]) \cap K_i^{[d_t]}([d_t] \rightarrow \overline{R(t)})$. In other words, $[p] \notin Rat(t)$. In order to see this it is enough to notice that $[d_t] \notin$ R(t). In fact, in this sub-case we have the following constraint for payoffs:

$$h_i([\beta(d_t)]) < h_i([\beta(a_t)]) = \beta_i(\mathrm{t}) = x$$

This (together with assumed properties of the transformation function t) is enough to guarantee that $[p] = [d_t] \in [(\overline{[t]} \cup [\eta_i < \mathrm{x}]) \cap K_i^{[a_t]}([a_t] \rightarrow [\eta_i \geq \mathrm{x}])$. And this, in turn, establishes that $[d_t] \notin \mathrm{R(t)}$.

Second sub-case: $[\beta_i(t)] = [d_t]$. In order to solve this sub-case it is enough to show that $[p] = [d_t] \notin Rat(v)$ – against the assumption that $[d_t] \in Rat$. An argument similar to the one given above suffices (by showing $[d_t] \notin R(v)$). ■

6. Discussion

The main proof in the previous section shows that a minimal extension of standard information structures provides a powerful tool to analyze games of perfect information. Conditional propositions are explicitly represented in the model. We use them to encode (as propositions) both the notion of rationality at a node and the notion of rationality as a disposition. The first notion is essentially Samet's. The second notion makes possible to formalize the idea that if all agents have the disposition to act rationally at all nodes, then they implement the BI solution in centipede games.

There is no need to posit that rationality is commonly known or believed. Or, following Samet's formulation, there is no need to assume that there is a common hypothesis of rationality among agents.

At first sight Theorem 5.1 has an existential character. It shows that there exist t-EIS models inducing the BI solution. Our main result can easily be presented, nevertheless, in a different form. In fact, the intended model used in the main result uses what we can call a *normal rationality condition*, by which the transformation function used in the encoding of rationality selects backward induction solutions in sub-games. This is a simple and natural selection function (saliently used in the informal analysis of cetipedes). A by-product of the main result is that normal rationality of the players is both a necessary and sufficient condition for the BI solution in *every* model of the game.

The demands on normal rationality that our transformation functions impose might be seen as substantial. Yet those constraints are no less demanding in alternative models which, in addition, impose considerable burdens on the amount of shared and commonly acquired knowledge needed to solve games. In our model the knowledge assumptions are kept to a minimum.

What would have happened had node H been reached? An agent can always ask this question (ex ante) in our model. The axioms we impose on transformation functions do not require that the agent should find the backwards induction solution at H every time he or she answers the question. Our axioms let the agent select other solutions or even suspend judgment upon supposing H. Some criticism of Samet's model has been in part based on the fact that it precludes these degrees of freedom — see [Halpern (1998)] and [Arlo-Costa and Bicchieri (1998)]. It may seem, however, that our main result depends on additional restrictions on transformation functions, which are not required axiomatically. Some obvious answers can be given. The first is that although our general result uses a special transformation function, many interesting games can be solved by permissible functions that allow agents to be endowed with considerably weaker cognitive powers. As an example, consider a centipede where option w_1 is down at root for player I and with payoffs $(1, 0)$, and where options w_2, w_3 and w_4 are, respectively, down for agent II with payoffs $(0, 2)$, down for agent I with payoffs $(3, 1)$ and, finally, across for player I with payoffs $(2, 3)$. In this game we have that *Rat* $= \{w_1\}$, even when one uses transformation functions *requiring* hypothetical suspension of judgement in the case of hypotheses about unreached nodes. In other words, $t_i(\{\mathrm{w}\}, \mathrm{H}) = \mathrm{H}$, whenever $w \notin \mathrm{H}$.

Answering the question: 'What would have happened had node H been reached?' depends on background knowledge about the payoff structure and on hypotheses about the rationality of other agents at different nodes. But even if all agents have the disposition to be rational, and each agent knows that, this does not necessarily force transformation functions to select backward induction solutions. In fact, at states where *Rat* is true, $K_i^{[v]}([v] \rightarrow [\beta(\mathrm{v})])$ need not be true for all nodes v. For example, consider again the aforementioned game. $K_i^{[v]}([v] \rightarrow [\beta(\mathrm{v})])$ is only true at w_3 if transformation functions are such that $t_i(\{\mathrm{w}\}, \mathrm{H}) = \mathrm{H}$, whenever $w \notin \mathrm{H}$.

So, on the one hand we can have games where *Rat* induces the backward induction solution, even though the transformation functions fail to select backward induction solutions at unreached nodes. On the other hand, normatively (or experimentally) justified transformation functions of this type might produce models where the disposition to act rationally at all nodes need not entail the BI solution. Models of this type could be justifiable to the extent that the corresponding transformation functions are justifiable.

Much of the recent interest in the explicit representation of the epistemic structure of games has grown out of the awareness that the known results about refinements and equilibrium selection are model-relative. Aumann's result [Aumann (1995)], showing that the BI solution is entailed by common knowledge of rationality, for example, depends on several theoretical assumptions. Knowledge representation matters in making those assumptions. Constructing a model involves selecting certain propositional attitudes (knowledge, belief, certainty) as the right attitudes, as well as building particular models of those attitudes. The crucial attitude selected in this paper is a form of supposing embedded in transformation functions. Emphasizing this propositional attitude makes other aspects of traditional models less crucial. In particular, neither common knowledge nor common hypotheses of rationality need to be assumed in order to derive the BI solution.

The models presented here are applicable to games of perfect information. Modeling conditionals in games of imperfect information requires developing a more sophisticated theory of conditionals. Games of imperfect information typically require to make hypotheses that are naturally modeled via functions T(P, H), where P is the current epistemic state, H the hypothesis and P is a complex proposition (a set of nodes, for example) [Bicchieri and Schulte (1997)], [Bicchieri (1993)]. In a companion paper we argue that the best modeling of conditionals in these games is syntactic rather than propositional [Bicchieri and Antonelli (1993)].

Throughout this paper we focused on studying a viable model for using conditional propositions in games of perfect information. The account that thus arises articulates a mathematical model of supposing. As we just argued, a more encompassing account of supposing might need a syntactic presentation. In addition, it should be stressed that the notion of supposing we are interested in is synchronic. Even when some of the ideas discussed here might be useful in order to model the (diachronic) notion of belief revision in games, this notion is more complex. Its careful study requires deploying concepts and formal apparatus which are beyond the scope of this paper. Our fundamental interest here just focuses on the interest and limits of the use of conditional propositions in modeling ('ex ante') supposing in games of perfect information.

References

[Adams (1975)] Adams, E., *The logic of conditionals*, Reidel, Dordrecht, 1975.

[Arlo-Costa (1999)] Arló-Costa, H., 'Belief revision conditionals: *basic* iterated systems,' *Annals of Pure and Applied Logic* 96 (1999), 3–28.

[Arlo-Costa and Bicchieri (1998)] ARLÓ-COSTA, H., and C. BICCHIERI, 'Games and Conditionals,' TARK VII, Theoretical Aspects of Rationality and Knowledge, July 22-24, 1998, Evanston, Illinois, USA, 1998.

[Arlo-Costa and Thomason (2001)] ARLÓ-COSTA, H., and R. THOMASON, 'Iterative probability kinematics,' *Journal of Philosophical Logic* 30 (2001), 479–524.

[Arlo-Costa (2001)] ARLÓ-COSTA, H., 'Bayesian Epistemology and Epistemic Conditionals: On the Status of the Export-Import Laws,' *Journal of Philosophy* Vol. XCVIII, 11 (2001), 555–598.

[Aumann (1995)] AUMANN, R., 'Backward induction and common knowledge of rationality,' *Games and Economic Behavior* 8 (1995), 6–10.

[Battigalli and Bonnano (1997)] BATTIGALLI, P., and G. BONANNO, 'The logic of belief persistence,' *Economics and Philosophy* 13 (1997), 39–59.

[Bicchieri and Schulte (1997)] BICCHIERI, C., and O. SCHULTE, 'Common Reasoning about Admissibility,' *Erkenntnis* 45 (1997).

[Bicchieri and Antonelli (1993)] BICCHIERI, C., and A. ANTONELLI, 'Game-theoretic Axioms for Local Rationality and Bounded Knowledge,' *Journal of Logic, Language and Information* 4 (1997).

[Bicchieri (1994)] BICCHIERI, C., 'Counterfactuals, Belief Changes, and Equilibrium Refinements', *Philosophical Topics* 21 (1994).

[Bicchieri (1989)] BICCHIERI, C., 'Self-Refuting Theories of Strategic Interaction: A Paradox of Common Knowledge ', *Erkenntnis* 30 (1989), 69–85.

[Bicchieri (1988)] BICCHIERI, C., 'Strategic Behavior and Counterfactuals', *Synthese* 76 (1988), 135–169.

[Bicchieri (1988a)] BICCHIERI, C., 'Common Knowledge and Backward Induction: A Solution to the Paradox,' in M. Vardi (ed.), *Theoretical Aspects of Reasoning about Knowledge*. Morgan Kaufmann Publishers, Los Altos, 1988.

[Bicchieri (1993)] BICCHIERI, C., *Rationality and Coordination*, New York: Cambridge University Press, 1993.

[Collins (1999)] COLLINS, J., 'Newcomb's Problem,' 1999.
http://www.columbia.edu/ jdc9/download.html

[Cross and Nute (1998)] CROSS, C., and D. NUTE, 'Conditional Logic,' in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, second edition*, Vol III Extensions of Classical Logic, Reidel, Dordrecht, 1998.

[Binmore (1996)] BINMORE, K., 'A note on Backward Induction,' *Games and Economic Behavior* 17 (1996), 135–37.

[Gardenfors (1978)] GÄRDENFORS, P., 'Conditionals and changes of belief,' in I. Niiniluoto and R. Tuomela (eds.), *The Logic and Epistemology of Scientific Change*, *Acta Philosophica Fennica* 30 (1978), 381–404.

[Gardenfors (1986)] GÄRDENFORS, P., 'Belief revisions and the Ramsey test for conditionals,' *Philosphical Review* 95 (1986), 81–93.

[Gardenfors (1988)] GÄRDENFORS, P., *Knowledge in Flux*, Bradford Book, MIT Press, Cambridge, Mass, 1988.

[Halpern (1998)] HALPERN, J., 'Hypothetical knowledge and counterfactual reasoning,' TARK VII, Theoretical Aspects of Rationality and Knowledge, July 83–97, 1998, Evanston, Illinois, USA, 1998.

[Hansson (1992)] Hansson, S. O., 'In Defense of the Ramsey Test,' *Journal of Philosophy* 89 (1992), 522–540.

[Joyce (199)] Joyce, J., *The Foundations of Causal Decision Theory*, Cambridge University Press, Cambridge, 1999.

[Levi (1988)] Levi, I., 'Iteration of conditionals and the Ramsey test,' *Synthese* 76 (1988), 49–81.

[Lewis (1973)] Lewis, D., *Counterfactuals*, Basil Blackwell, Oxford, 1973.

[Lewis (1976)] Lewis, D., 'Probabilities of conditionals and conditional probabilities,' *Philosophical Review* 85 (1976), 297–315.

[Ramsey (1990)] Ramsey, F. P., *Philosophical Papers*, H. A. Mellor (ed.), Cambridge University Press, Cambridge, 1990.

[Samet (1996)] Samet, D., 'Hypothetical knowledge and games with perfect information,' *Games and economic behavior* 17 (1996), 230–251.

[Stalnaker (1968)] Stalnaker, R. C., 'A theory of conditionals,' *Studies in Logical Theory, American Philosophical Quarterly* Monograph:2 (1968), 98–112.

[Stalnaker and Thomason (1970)] Stalnaker, R. C., and R. Thomason, 'A semantical analysis of conditional logic,' *Theoria* 36 (1970), 246–281.

[Rubinstein (1997)] Rubinstein, A., *Modeling Bounded Rationality*, MIT Press, 1997.

Horacio Arló-Costa
Department of Philosophy
Carnegie-Mellon University
Pittsburgh, PA, USA
hcosta@andrew.cmu.edu

Cristina Bicchieri
University of Pennsylvania
433 Logan Hall,
Philadelphia, PA 19104
cb36@sas.upenn.edu