# Tracking truth: Knowledge and conditionals in the context of branching time


Horacio Arló Costa
CMU
hcosta@andrew.cmu.edu

Rohit Parikh
CUNY
RParikh@gc.cuny.edu


June 1, 2006

Van Benthem has recently remarked that Nozick's notion of knowledge via 'tracking the truth' constitutes '[...] an open challenge' to logicians. It is easy to see that its adoption will block laws such as Distribution or Introspection. But are there any valid inference patterns left? Given some plausible background logic of belief and counterfactuals, what is the complete set of validities of Nozick's K? [35].' In a recent paper Arló-Costa and Parikh [1] tackled this problem and offered a parametrical account of the resulting logics of knowledge by considering various such plausible background conditional and doxastic logics.

The logics of knowledge that thus arise are all non-normal modal logics, where some aspects of logical omniscience fail [29] and which we modeled by appealing to neighborhood semantics [2]. So the standard relational models inspired by the seminal work of Kripke are inapplicable here. In fact the knowledge operator is not closed under entailment and therefore there is no known technique to formulate a normal logic that can simulate it [19]. Nevertheless other important axioms, like Adjunction, are obeyed.

It is important to emphasize here that this is a case where logical omniscience fails for normative rather than descriptive reasons (like lack of computational power). Kyburg and associates [20], [21] also offered normative reasons of different kind for abandoning closure. In this case it is Adjunction that fails, while other forms of logical omniscience, like closure under entailment, are obeyed.

Logical closure is neither intrinsically good nor bad as a normative epistemic or doxastic property. All depends on the notion that is being modeled. For example, if one focuses on modeling *commitments to full belief* (as Levi does in [24]) then closure is appropriate. Our doxastic models below can be understood in this way.

In this paper we tackle the problem of modeling the underlying conditional logic in Nozick's notion of knowledge and by offering a complete characterization of the notion of knowledge as tracking truth that arises from plausible minimal assumptions about belief and conditionals. More importantly, following the most sophisticated presentation of Nozick's theory, we formalize explicitly the notion of *method* of inquiry. This goes to the heart of Nozick's proposal and further philosophical reflection about it in the works of, for example, Williamson [36] and DeRose [12].

The lack of closure in Nozick's notion of knowledge arises normatively even when both the underlying doxastic and conditional notions are normal operators satisfying all aspects of logical omniscience. But the logical framework we offer here is rather flexible and nothing precludes studying the notions of tracking knowledge arising from weaker doxastic and conditional models where various aspects of logical omniscience fail either normatively (for example, when conditionals are constructed probabilistically, as we explain in the last section) or descriptively.

# 1 Knowledge as tracking truth

The problem of how to represent the knowledge of interactive agents has been extensively studied during recent years by philosophers, computer scientists, economists and logicians. Some of the most recent models are quite rich and sophisticated. They take into account not only the fact that knowledge is social but also the fact that representing the knowledge of a single agent requires modeling other attitudes as well, like belief and conditional belief. So, some of the most recent formal models of knowledge are built in multi-agent and multi-modal epistemic environments.

Many informal characterizations of knowledge recently offered by philosophers are indeed of this type. An example of a notion of knowledge clearly grounded in a modal environment is Robert Nozick's notion of knowledge as *tracking the truth.* Like many philosophers since at least Plato, Nozick proposes to formulate further conditions for attributing knowledge of $A$ to an agent $S$ (where A is a sentence in a target language) that go alongside:

**(1)** $A$ is true

**(2)** $S$ believes $A$

The idea is that each further condition has to be necessary for knowledge, so any case that fails to satisfy it will not be an instance of knowledge. In [28] (pages 172-176) Nozick proposed the following two additional conditions:

**(3)** If $A$ weren't true, S wouldn't believe that $A$

**(4)** If $A$ were true, S would believe that $A$

The conditionals are supposed to capture the way in which the agent's beliefs are *sensitive* to the truth-value of $A$. (3) tells us how the agent's belief state is sensitive to $A$'s falsity and (4) tells us how the agent's belief state is sensitive to $A$'s truth.

Nozick's article questioned some of the best existing models of epistemic operators and conditionals. In particular, Nozick had to challenge the 'official' model of counterfactual conditionals offered by David Lewis in [25]. In fact, according to this model condition (4) is redundant. In Lewis's account the truths of antecedent and consequent of a conditional are jointly sufficient for the truth of the conditional. Nozick argued as follows: the coin was tossed and it landed heads, from this does not follow that it if were tossed it would land heads. Motivated by what he saw as a deficiency in Lewis's account, Nozick formulated an alternative semantics for conditionals in his book. Nozick's remarks, however, are made in passing, and gathered in a footnote.

Nozick was aware of the fact that the possible-worlds semantics used by Lewis was controversial, but he tried to formulate his critical remarks within this framework. Since then many philosophical logicians and epistemologists have followed the same path. Many of the best current criticisms and presentations of Nozick's account are today equally guarded when it comes to buy into all the details of Lewis's semantic program for modalities and conditionals, but also equally disposed to introduce modifications that will make it suitable for epistemological applications [36] [12].

We think that a formal model of Nozick's notion of knowledge requires more drastic departures, not only from some of the semantics utilized in possible-worlds models of conditionals, but also from the Kripkean semantics that has been usually employed in order to formalize knowledge and belief both in philosophical logic and in computer science [14], [17]. As we explained above the notion of knowledge proposed by Nozick fails to respect logical closure and therefore there is no Kripkean modal system that can represent it.

We are primarily interested in developing formal tools useful for representing and discussing Nozick's insights about knowledge. As the reader will soon see, this will also require rethinking some important foundational problems like: what counts as a possible world in a multi-agent and a multi-modal representation of epistemic attitudes? It will also lead to a thorough utilization of foundational ideas in modal logic proposed by Dana Scott and by Richard Montague in the early 1970's. So, even when one of the byproducts of our analysis will be to produce an axiomatization of Nozick's notion of knowledge as well as a model of the notions of belief and conditional belief that are needed to capture it, we centrally intend to use Nozick's notion of knowledge as a motivating example capable of showing some of the weak points of many of the formal tools commonly used in formal epistemology. We also intend to develop concrete alternatives that we think have a wide area of applicability in fields that go from theoretical computer science and mathematical economics to applied logic and epistemology per se.

# 2 Some basic syntactic constraints on conditionals

One possible strategy here is to take Nozick's definition for granted and then focus on the logical structure of the epistemic logic that arises when the underlying conditional logic is constrained by axioms sufficiently strong to reflect syntactically the notion of conditionality pre-systematically involved in Nozick's account of knowledge.

Although Nozick protested against Lewis's possible worlds semantics, he did make crucial use of it and apparently he never questioned central assumptions made by Lewis, like the assumption that conditionals (of the type that interest him) carry truth-values. The following conditions therefore seem salient and required for these type of conditionals.

$$\text{(R)} \quad A > A \qquad \text{(LLE)} \quad \frac{\vdash A \leftrightarrow B,\ A > C}{B > C}$$

$$\text{(RW)} \quad \frac{\vdash A \rightarrow B,\ C > A}{C > B} \qquad \text{(CMP)} \quad \frac{A > B}{A \rightarrow B}$$

$$\text{(AND)} \quad \frac{A > B,\ A > C}{A > B \wedge C} \qquad \text{(OR)} \quad \frac{A > C,\ B > C}{A \vee B > C}$$

$R$ (sometimes called Reflexivity) seems constitutive of any notion of supposition (either ontic or epistemic). Moreover, failures of this axiom might lead to failures in attributing knowledge even in cases where the agent only has reliable and fully justified true beliefs (in the presence of Nozick's definition). Conditional Modus Ponens ($CMP$) seems constitutive of most reasonable notions of the counterfactual conditional and of many epistemic notions as long as the letters $A, B$ stand only for purely Boolean sentences – see [6] and [31] for failures of $CMP$ in epistemic and probabilistic models.

The notation we are adopting here is the one usually employed in recent work on non-monotonic logic (see [23] for example). Left Logical Equivalence $LLE$ – also known as $RCEA$ in the literature of modal logic (see [9], page 269) – is a very weak rule of inference, which presupposes that we are assuming a principle of 'irrelevance of syntax' in our conditional logic. This is so as long as we also assume a similar rule for consequents as well (which we do via (RW)). This additional rule allowing for substitution of equivalent sentences in consequents is usually called $RCEC$. In general, a conditional logic closed under $LLE$ and $RCEC$. is called *classical*.

We assumed here a stronger rule: Right weakening $RW$, also known as $RCM$. Obviously, the resulting logic is also classical (every logic closed under $LLE$, which is in addition to closure under $RW$ is also closed under $RCEC$).

In addition we have the AND and OR conditions, which deal with the corresponding connectives. Actually we will focus here on an underlying conditional logic that only adds OR to the axioms $R$ and $CMP$ and the rules $LLE$ and $RW$. The resulting conditional logic is still rather weak. Most conditional logics impose *regularity* conditions:

$$\text{(RCR)} \quad \frac{B \wedge B' \to C}{(A > B) \wedge (A > B') \to (A > C)}$$

$$\text{(RCK)} \quad \frac{(B_1 \wedge ... \wedge B_n) \to B}{((A > B_1) \wedge ... \wedge (A > B_n)) \to (A > B)}$$

Conditional logics closed under $LLE$ and $RCK$ are usually called *normal*. They admit semantics in terms of selection functions, usual in the field. The logic we are assuming here is not regular (it is not closed under $RCR$) and therefore it is now normal either. Semantics for this type of logics have been proposed by Chellas and others in terms of dyadic extensions of the semantics for (monadic) modalities by Dana Scott [33] and Richard Montague [27].

We will not focus on semantic issues regarding the underlying conditional logic until later. We can point out here in passing, nevertheless, that failures of normality can appear both at the level of the monadic and dyadic operators, and in both cases it can have similar kind of semantic treatments.

# 3 The doxastic core

We need to focus here in the underlying notion of belief. As in the previous case we will present the main conditions for a doxastic operator, indexed for an agent $i$ and a time $t$.

### Axioms for the doxastic core

**(DCon)** $\neg B_i^t(\mathit{false})$

**(DC)** $(B_i^t(A) \wedge B_i^t(B)) \to B_i^t(A \wedge B)$

**(DM)** $B_i^t(A \wedge B) \to (B_i^t(A) \wedge B_i^t(B))$

**(DN)** $B_i^t(\mathit{true})$

### Rules of inference for the doxastic core

**(RDE)** From $(A \leftrightarrow A')$ infer $(B_i^t A \leftrightarrow B_i^t A')$

This system is also rather weak, but (in view of (DM)) it is a normal system. It is actually the system **K** plus the consistency condition DCon. This is the bare minimum we need for our first characterization of knowledge, as we will show in the next section.

### 3.1 Other doxastic notions

Some recent work in philosophy focusing on epistemic notions related to Nozick's tracking notion, have studied probabilist variants of the standard definition offered by Nozick. One example is [32] where high conditional probability is used rather than tracking conditionals in order to study a notion akin to Nozick's. Although the author does not give much detail about the underlying doxastic notion, it seems that symmetry would require treating unconditional belief probabilistically as well in terms of high (unconditional) probability.

In that case the axioms given above would be too strong. In particular the axioms *DC* would fail. We do not here replace Nozick's conditionals by conditional probability (not even if this involves a non-Kolmogorovian notion of conditional probability). Given well-known facts in conditional logic, this seems tantamount to a change of theme, although we do not want to enter into this issue here in detail.

Since we do not replace conditionals by conditional beliefs encoded probabilistically as high conditional probability, we do not treat belief probabilistically either. We propose instead to have a weak but normal notion of belief in the doxastic base obeying all forms of logical omniscience. Of course, nothing in our setting precludes representing formally (in addition) probabilitic beliefs or epistemic conditionals of this kind. Probabilistic beliefs and conditional beliefs of this kind might be used to keep track of the conditionals that agents *accept*, not necessarily in the sense of accepting them as true, alongside other ontic conditionals used in epistemic attributions of the kind that seemed to have attracted Nozick's interest. The last section offers a sketch of a model of this kind for conditionals.

## 4 Knowledge as tracking true: axioms

The doxastic core plus the minimal constraints on conditionals presented above provide sufficient conditions for the following syntactic characterization of knowledge.

**Axioms for Knowledge**

**(KT)** $K_i^t(A) \rightarrow A$

**(KN)** $K_i^t(\mathbf{true})$

**(KCon)** $\neg K_i^t(\mathbf{false})$

**Rules of inference for Knowledge**

**(RKE)** From $A \leftrightarrow B$, infer $K_i^t(A) \leftrightarrow K_i^t(B)$,

**(RKC)** From $(K_i^t(A) \wedge K_i^t(B)$, infer $K_i^t(A \wedge B)$

We provide here an explicit argument for $RKC$ . Assume that $(K_i^t(A) \wedge K_i^t(B))$. We need to prove that $K_i^t(A \wedge B)$. The first condition we need to establish is that:

$(A \wedge B) > B_i^t(A \wedge B)$

Consider the following instance of $R$:

$(A \wedge B_i^t(A) \wedge B \wedge B_i^t(B)) > (A \wedge B_i^t(A) \wedge B \wedge B_i^t(B))$

By the main assumption and $CMP$ we can detach $A \rightarrow B_i^t(A)$ (and similarly for $B$). Therefore by $LLE$ we have that:

$(A \wedge B) > (A \wedge B_i^t(A) \wedge B \wedge B_i^t(B))$

And then by $RW$ and $DC$ we are done. Concerning $\neg(A \wedge B) > \neg B_i^t(A \wedge B)$, the argument follows immediately by appealing to $RW$, $DM$, and $OR$.

Now this argument does not give us

**(KC)** $(K_i^t(A) \wedge K_i^t(B) \rightarrow K_i^t(A \wedge B)$

Now, if we strengthen our conditional base we can get a condition approximating $KC$. We can add *cautious monotony* $CM$ establishing that from $A > C$ and $A > B$ we can derive $A \wedge B > C$. Let now introduce the notation $A <> B$ to abbreviate that both: $A > B$ and $B > A$. Then we can have an axiom indicating that if $A$ and $B$ are counterfactually (or causally) correlated then if both $A$ and $B$ are known their conjunction is also known:

**(CKC)** $A <> B \rightarrow [(K_i^t(A) \wedge K_i^t(B) \rightarrow K_i^t(A \wedge B)]$

This is the closest to the pure $KC$ condition we can get with this conditional base and doxastic core.

# 5 Knowledge as tracking true: first semantic characterization

Following the terminology introduced by Chellas in [9] we will call *classical* a system of modal logic closed under the rule of inference $RKE$ presented above (and where a dual of the main modality is introduced via the standard definition).

It is well known that many classical but non-normal systems of modal logic of the type that interest us here cannot be modeled via standard (Kripkean) relational semantics. Some of these systems can be simulated by normal mono-normal or poly-normal logics [19] but this applies only to non-normal classical systems which are monotonic, and monotony (closure under entailment of the knowledge operator) fails in the case that interest us.

So, let $\Sigma$ be a classical system of modal logic and let $|A|_\Sigma$ be the set of maximal and consistent extensions of the underlying language containing $A$ – we will assume that this language contains countaby many atomic propositions $p_1, ..., p_n, ...$. We will call $|A|_\Sigma$ $A$'s *proof set*.

Let, in addition, $\langle W, E, V \rangle$ be an epistemic neighborhood model where $W$ is a set of points, $E$ is a function mapping points in $W$ to sets of propositions, and where $V$ is a valuation. We will be interested later on in applications where the points in $W$ are interpreted as pairs (branch, times) in a branching time structure. We will make explicit this connection below. For the moment we will focus instead on a presentation where the points in $W$ are unstructured primitives.

$\mathcal{M}$ is a *canonical epistemic neighborhood model* for a classical system $\Sigma$ if and only if:

**(1)** $W$ is the set of maximal and $\Sigma$-consistent extensions of the underlying language.

**(2)** For every $w$ in $W$, $K_i^t(A) \in w$ if and only if $|A|_\Sigma \in E(w)$

**(3)** $V(p_n) = |p_n|_\Sigma$, for $n = 0, 1, 2, ..$

If $\mathcal{M} = \langle W, E, V \rangle$ is an epistemic neighborhood model, $|\mathrm{A}|^{\mathcal{M}}$, as before, represents $A$'s truth-set. Then it is easy to show that if $\mathcal{M}$ is a canonical epistemic neighborhood model for a classical system $\Sigma$, $|\mathrm{A}|^{\mathcal{M}}$ if and only if $|A|_\Sigma$.

The simplest example of a canonical model for a system $\Sigma$ is that in which, for each point $w$ in the model, $E(w)$ is the set $\{|A|_\Sigma\colon K_i^t(A) \in w\}$. In other words, $E(w)$ consists just of those proof sets $|A|_\Sigma$ such that $K_i^t(A) \in w$ and nothing else. The *smallest* canonical model for $\Sigma$ has this form. The *largest* canonical model for $\Sigma$ is an extension of the smallest canonical plus every set $X$ in $W$ that is not a proof set in $\Sigma$ for any sentence. And, of course, if the neighborhoods of a model always contain the set $\{|A|_\Sigma\colon K_i^t(A) \in w\}$ plus a collection of non-proof sets for $\Sigma$, it will still be a canonical epistemic neighborhood model for $\Sigma$. This issue will be important below.

The system $E$ which is closed under $RE$ and that only contains the definition of the dual of the knowledge operator is completely characterized by the class of epistemic neighborhood models. We have in addition the following result:

**Theorem 5.1** *Let $\mathcal{M} = \langle W, E, V \rangle$ be the smallest canonical epistemic neighborhood model for a classical system $\Sigma$ containing:*

*(KT)* $K_i^t(A) \rightarrow A$

*(KN)* $K_i^t(\mathbf{true})$

*(KCon)* $\neg K_i^t(\mathbf{false})$

*(KC)* $(K_i^t(A) \wedge K_i^t(B) \rightarrow K_i^t(A \wedge B))$

*then for evey $w, X, Y$ the model $\mathcal{M}$ is respectively:*

*(Alethic) If $X \in E(w)$, then $w \in X$*

*(Contains the unit) $W \in E(w)$*

*(Consistent)* $\{\emptyset\} \notin E(w)$

*(Closed under intersections) If* $X \in E(w)$ *and* $Y \in E(w)$*, then* $X \cap Y \in E(w)$*,*

Therefore classical logics constrained by (Con), (C), (N), and (T) axioms are complete with respect to the class of consistent neighborhood models, closed under intersections, containing the unit and alethic. The smallest system of this class, namely EConCNT, is characterized by this class of models, since it is also sound with respect to it.

Now, it is of some interest to notice that if $\mathcal{M}$ is the smallest canonical epistemic neighborhood model for a classical system $\Sigma$, then its infinite closure is also a canonical epistemic neighborhood model for the system. A compactness argument can be presented in order to substantiate this claim. Here is the gist of the argument, which is based on showing that the infinite intersection of a neighborhood is a non-proof set for $\Sigma$. Suppose that there is a sentence $A$ such that:

$\bigcap\{|A|_\Sigma\colon K_i^t(A) \in w\} = |A|_\Sigma$

In other words, $A$ is exactly a member of every $\Sigma$-maximal set of sentences $v$ in $\bigcap \{|A|_\Sigma\colon K_i^t(A) \in w\}$. It is easy to see that:

$v \in \bigcap \{|A|_\Sigma\colon K_i^t(A) \in w\}$ if and only if $\{A\colon K_i^t(A) \in w\} \subseteq v$

By hypothesis $A \in v$ for every $\Sigma$-maximal set $v$ such that $\{A\colon K_i^t(A) \in w\} \subseteq v$. In other words $A$ belongs to every $\Sigma$-maximal extension of the set $\{A\colon K_i^t(A) \in w\}$. Therefore it is immediate by Lindembaum's lemma that $A$ is $\Sigma$-deducible form this set of sentences:

$\{A\colon K_i^t(A) \in w\} \vdash_\Sigma A$

By the compactness of $\vdash_\Sigma$ we know that there is a finite set of sentences $\{A_1, ..., A_n\} \subseteq \{A\colon K_i^t(A) \in w\}$ such that:

$\{A_1, ..., A_n\} \vdash_\Sigma A$

But, since the set $\{A_1, ..., A_n\}$ is finite, the monotony of $\vdash_\Sigma$ indicates that it is always possible to find a $\Sigma$-maximal set of sentences $v'$ not in $\bigcap \{|A|_\Sigma\colon K_i^t(A) \in w\}$ to which $A$ belongs. This completes the argument.

One immediate consequence of this argument is that EConCNT is also determined by the class consistent neighborhood models, closed under intersections, containing the unit and alethic and which also contain its infinite intersection.

*Augmented models* satisfy the following condition: (Augmentation). $X \in E(w)$ **iff** $\cap E(w) \subseteq X$. It is clear therefore that the augmentation of the smallest canonical is not a canonical, for an augmented model is automatically also closed under supersets.

# 6 Interactive Belief, Conditionals and the role of Possible Worlds

We will start in this section with a model of multi-agent belief and then we will make it more sophisticated in order to encompass conditionals as well. This will be a first step towards developing a semantics for the underlying conditional logic and the doxastic core. We will use a framework popular in multi-agent epistemic logic (mainly used in computer science, and of increasing popularity for other applications). The model in question is offered by Joseph Halpern [17] in order to produce a formal analysis of vagueness. Many of the notions utilized in the model as well as the general approach of which the model is an instance comes from the various models of knowledge presented in [14].

Let's consider a modal logic with $n$ belief operators, $B_1, ..., B_n$, one for each agent included in the model. So, $B_1 A$ is interpreted as 'agent $i$ believes $A$'. A *belief structure* $M$ has the form $\langle W, P_1, ..., P_n, \pi_1, ..., \pi_n \rangle$ where $P_i$ for $i$ = 1, ..., $n$, is a subset of the set of worlds $W$ and where $\pi_i$ is an interpretation that associates with each atomic proposition a subset of $W$. In many models each $P_i$ is interpreted as the set of worlds that the agent considers *plausible*.

Now, it is important to remark that in most models (certainly in the ones presented in [14] and [17]) it is assumed that $W \subseteq O \times S_1 \times ... \times S_n$ where $O$ is the set of *objective* states and $S_i$ is a set of *subjective* states for agent $i$. Therefore worlds have the form $(o, s_1, ..., s_n)$. In multi-agent systems $o$ is called the *environment state* and each $s_i$ is called a *local* state for the agent in question.

Halpern characterizes an agent's subjective state $s_i$ by saying that it represents '$i$'s perception of the world and everything else about the agent's make-up that determines the agent's reports'. David Lewis in one of his last published papers in epistemology [25] also appeals to hybrid and composite *possibilia* of this sort. Lewis clearly specifies that these possibilities cannot be either just 'real' possibilities or just 'epistemic' possibilities. He stipulates that '...they are not just possibilities as to how the whole world is; they include also possibilities as to which part of the world is oneself, and as to when it is now'. Lewis leaves as an open question whether the content of a perceptual experience $E$ (used to eliminate a possibility $w$) can be fully characterized propositionally. Halpern's models go well beyond this by taking not only the 'subjective' states in each $S_i$, but also the 'objective' states in $O$ as primitives for which there is no propositional representation. In fact, notice that if $o$ in $(o, s_1, ..., s_n)$ is represented via a proposition (understood as a set of worlds) or if each $s_1$ is represented by a set of propositions, then in order to avoid circularity the whole account needs to abandon the axiom of foundation in classical set theory (there are indeed sophisticated models of epistemic modalities of this type - see, for example, [8] and the discussion in [7]).

The issue of what counts as a possible world in this construction will be important in a moment when we add conditionals to the doxastic kernel of the model. But for the moment we will adopt it in an unmodified manner in order to see how it operates. For example an important issue will be how we can use possible worlds to interpret primitive or atomic propositions. Notice first that we can easily define some notions of 'indistinguishability'. One can write $w \sim_i w'$ to indicate that the possible worlds $w$ and $w'$ agree on $i$'s subjective state. One can also write $w \sim_o w'$ to indicate that the worlds $w$ and $w'$ agree on their objective parts.

With this technical definition we can see that if $w = (o, s_1, ..., s_n) \in \pi_i p$ for an atomic proposition $p$, then if $w \sim_i w'$ and $w \sim_o w'$ then $w' \in \pi_i p$. In other words, the evaluation of the truth of atomic propositions does not only depend on the objective part of the state in question but also on the subjective part - which means that even the most elementary truth evaluations have a strong hidden epistemic content.

The notation $(M, w, i) \models A$ indicates that $A$ is true according to agent $i$ at world $w$. We already know that $(M, w, i) \models p$ iff $w \in \pi_i p$. Then $\models$ can be defined for other formulas by induction. For example, we have that $(M, w, i) \models \neg A$ iff $(M, w, i) \not\models A$. Notice that this simple definition will immediately have epistemic repercussions, for it entails that for every primitive proposition each agent is prepared to say whether or not the proposition holds. This is clearly unrealistic and leads to the addition of epicycles to the construction, which we prefer to ignore for the moment. Let's focus instead on the semantics of the belief operators.

The idea is that $(M, w, i) \models B_i A$ holds if and only if $(M, w', i) \models A$ for all $w'$ such that $w \sim_i w'$ and such that $w' \in P_i$. In other words $B_i A$ is true if $A$ is true at all plausible states that agent $i$ considers possible. It is easy to see that if we define a relation $\mathcal{R}_i$ on worlds consisting of all pairs $(w, w')$ such that $w \sim_i w'$ and such that $w' \in P_i$, the defined relation is Euclidean and transitive. This in turn determines a belief operator satisfying the axioms of the modal system $KD45$.

As we explained above this is a relatively popular way of constructing a basic doxastic model in a multi-agent system. We presented some criticisms to this methodology above, but, as we also explained above, we will make the model more sophisticated (i.e. we will transform it in a multi-modal model as well) before giving a more complete evaluation and an alternative. It should also be said before proceeding further that even if the model is not a multi-agent model many philosophers have assumed some of the basic tenets assumed above about the encoding of possible worlds in their informal accounts of single-agent belief and knowledge.

Now, we need to extend the analysis in order to model conditionals. As we explained above, Nozick does not endorse an unmodified form of Lewis's analysis of conditionals. Nevertheless the departure from Lewis' analysis is not drastic and it can be easily obtained by reformulating Lewis' definition in a particular case.

Without providing details this is the basis of Lewis' analysis. A conditional $A > B$ is true at a possible world $w$ if and only if either $A$ is true at no possible world, or, for at least one possible world $w'$, $A$ is true at $w'$ and $B$ is true at every possible world at least as close in the relevant aspects as $w'$ is to $w$. Nozick cannot accept this for it requires that $A \wedge B$ entails $A > B$. So, he reformulates it as follows: If $A$ is true at $w$, then $A > B$ is true at $w$ if and only if $B$ is true at every world close to $w$ at which $A$ is true. Lewis in his book *Counterfactuals* presents axiomatically various systems of conditionals (including conditionals that are not counterfactuals) aside from his 'official' system **VC**. One of these systems (**VW**) is based on the same idea that motivates Nozick's reformulation.

It should be remarked though that the notion of similarity involved in Nozick's reformulation is not (cannot be) Lewis' notion of *comparative similarity* among worlds, which implements a mereological notion of distance between worlds. As Williamson remarks the notion is epistemically motivated (and it is also reminiscent to the notion presented in [36], section 5.3).

Lewis' notion of comparative similarity among 'standard' possible worlds (in the traditional sense of possible worlds understood as primitives without parts or components) is not problem-free (see [15]). But how should we extend or modify the notion of comparative similarity when the compared possible worlds are the composite entities $(o, s_1, ..., s_n)$ that we considered above? Should we say that two possibilities sharing exactly the evidence of a group of agents resemble each other even when in other respects they might be very dissimilar? David Lewis considered this question in a recent paper on epistemology and answered this question negatively. Nevertheless, he recognized that his answer was *ad hoc* and that he did not know how to tackle this issue in a principled way. When we consider a group of agents things get worse rather than better. Should we consider that two possibilities resemble each other sufficiently if they share the evidence of a sufficiently large group of agents and are in other respects not very dissimilar? Even for the single agent case, it is not clear how we should proceed in order to determine doxastic similarity, even when worlds are sufficiently similar with respect to other non-doxastic components.

Consider the following example that touches directly the issue of sensitivity. Consider two worlds $w = (o, s_1)$ and $w' = (o, s'_1)$. Say that a factual sentence $A$ is true at both worlds and assume as well that while we have $(M, w, 1) \models \neg B_1 A$, we also have $(M, w', 1) \models B_1 A$. In other words, the model represents the case where the agent comes to believe $A$ and he does so correctly - he discovers the truth of a new proposition.

Now let's evaluate $A > B_i A$ at $(M, w', i)$. Is it the case that: $(M, w', i)$ $\models A > B_i A$? All depends on whether $w$ resembles $w'$ sufficiently. We assumed that the objective parts of $w$ and $w'$ are identical. And we can assume as well that if there is any workable notion of doxastic similarity $w$ is indeed sufficiently close to $w'$ – we might assume that the only salient belief that separate these states is the belief regarding $A$. But then we have that $(M, w', i) \not\models A > B_i A$. Which determines that when an agent correctly comes to believe a new proposition he never gains knowledge of this proposition. It seems easy to extend this result for groups of agents, in such a way that it gets even more poignant.

The notion of resemblance or similarity among 'composite' worlds of the kind we are studying here is even more problematic than the notions of similarity based on mereological considerations (which could be more appropriate for ontological applications). If the latter notion is poorly understood or if you happen to think that it poses hard problems for the possible worlds account of conditionals, things get considerably worse when the worlds are suitably enriched in order to apply them in epistemology. You have the basic problems that you have in the ontological case and more.

# 7 An alternative way of building a doxastic model

There is an alternative manner of developing epistemic logics in a multi-agent and multi modal environment. The basic idea is to develop an epistemic application for the semantic framework commonly called 'neighborhood' semantics. This kind of semantics was advocated by Dana Scott [33] and by Richard Montague [27] as an alternative and a generalization of the standard Kripkean account of modalities. We already have made use of it above to fully characterize a first version of knowledge as tracking truth.

As we have seen most of the existing accounts, developed either by computer scientists or by philosophers, have opted for making a crucial foundational modification in the standard account of modalities where the possible worlds are no longer treated as primitives but rather as composite entities with doxastic components. This strategy has also been advocated to some extent by game theorists interested in developing epistemic justifications for certain kind of games, [7].

In epistemic logic there is a genuine need for a representation capable of keeping track of the epistemic state of the agent(s) that inhabit this or that world. But, as we discussed above, it seems that as long as such an epistemic state is encoded propositionally this type of representation cannot be taken at face value. Or it can be taken at face value only by giving up some important axioms of set theory.

Worlds can be construed as composite entities encoding the *local* states of agents together with the *environment* state as long as neither representation is propositional. For example, in multi-agent systems the local states might be encoded as sets of actions taken by agents at some point in time, or messages passed from agent to agent, or pairs of truth assignments and sets of formulas, etc.[1] But, local states cannot be encoded as propositions (or

[1] See [14] section 4.4 for several examples.

sets of propositions) without circularity, as long as we construe (as usual) propositions as sets of worlds themselves.

There is yet another alternative to which we have appealed various times above. The idea is to associate a *neighborhood* of propositions to each world. The intuition is that the set of propositions in the neighborhood of world $w$ is the set of propositions accepted or believed by certain agent at $w$. Now the propositions in question are no longer *part* of $w$, they are *associated* with $w$. This is enough to break the circularity. At the same time the solution is expressive enough to keep track of the (propositional representations of) epistemic states of agents at worlds.

## 7.1 Branching time structures

There is an additional and different way of having worlds endowed with some structure. This alternative way is made possible by representing time explicitly in the models, a strategy usually employed in learning theory [18]. Here is the basic idea:

Our basic temporal structure is the one utilized in CTL (Computation Tree Logic as presented in [22] - see also [10]). Let $R$ be a transition relation on the set of states $W$ that we assume is serial (i.e. for every $w \in W$ there exists $w' \in W$ such that $\langle w, w' \rangle \in R$). Now we need in addition a distinguished state $w^0$, called the initial state. We also need a labeling function $L : W \rightarrow 2^{AT}$ mapping each state to a set of atomic propositions in the set $AT$ of atomic propositions of the underlying language. The atomic propositions in question should be true in that state. This labeling function give us a finitary (but faithful) representation of the *environment* at states. The representation is propositional.

So, we have a temporal structure $S = \langle W, R, w^0, L \rangle$. A *path* in $S$ is a sequence of states $p = w_0, w_1, \ldots$ such that for every $i \geq 0$, we have $\langle w_i, w_{i+1} \rangle \in R$. We use $p^i$ to indicate the *suffix* of $p$ starting at $w_i$.

Let $\Pi_{w^0}$ the set of root paths starting at the initial state. As before we use $\pi^i$ to indicate the *suffix* of the root path $\pi$. Notice that suffixes can now be interpreted as well as *instants*.

CTL is basically a branching time structure. Now, if $I$ is the set of instants and $\Pi_{w^0}$ the set of permissible roots paths we add an epistemic component to this structure by simply adding a a function $E : I \times \Pi_{w^0} \rightarrow 2^{2^{\Pi_{w^0}}}$. So now we have more complex structures $K = \langle W, R, w^0, L, E, \Lambda \rangle$, where $\Lambda$ is the set of all permissible iterated learning sequences (see the second part of [5] for a review of qualitative methods if iterated change and [18] for an alternative account in terms of learning theory). See [30] also for an alternative presentation of a temporal structure which is useful in this setting.

Notice that neighborhoods are assigned to pairs of instants and root paths, rather than to worlds taken as unstructured primitives, so we can have that identical factual states are assigned different neighborhoods. The epistemic state at a world is not supervenient on the facts that hold true at this structured point but a function of evidence received with each tick of the clock and the initial background knowledge. Notice that now the worlds in our model have been given a certain amount of structure (although this structure is rather different than in other well-knonw accounts [14], which are unworkable in this setting where we are dealing with non-normal logics). And, as Thomason and Gupta pointed out in [34] this added structure may provide some purchase on the notion of world similarity that is crucial to interpret conditionals. So the increase of expressive power gives us nice insights on the most intractable aspect of interpreting conditionals, i.e. the epistemic and ontological origin of world similarity. One constraint that we borrow from Thomason and Gupta is to require the principle that they call *Past Predominance*. The idea here is that in determining how close $\pi^t$ is to $\pi'^t$ past closeness predominates over future closeness (i.e. the portions of $\pi$ and $\pi'$ not after $t$ predominate over the rest of $\pi$ and $\pi'$).

The interaction of tense and conditionals is a complex affair. We will reduce part of this complexity in this presentation by evaluating only conditionals of the sort that interest us, namely tracking conditionals involving only combinations of Boolean and doxastic formulas, but not additional temporal quantifiers and operators.

**DEFINITION 7.1** $\mathcal{M}_\lambda = \langle S, E, \Lambda \rangle$ *is a learning frame if and only if:*

**(1)** S is a CTL structure $\langle W, R, w^0, L \rangle$

**(2)** $E$: $I \times \Pi_{w^0} \rightarrow 2^{2^{\Pi_{w^0}}}$

**(3)** $C$: $(I \times \Pi_{w^0} \times 2^{\Pi_{w^0}}) \rightarrow 2^{2^{\Pi_{w^0}}}$

**(4)** $\Lambda$: $\Pi_{w^0} \rightarrow E$

The following definition makes precise some salient aspects of the notion of truth in a model (we do not enter into interactions of tense, conditionals and belief, as we explained above):

**DEFINITION 7.2** Truth in a neighborhood model*: Let $\pi^i = w_i$ be a state in a model $\mathcal{M}_\lambda = \langle S, E, \Lambda, \models \rangle$. The following clauses are added in order to determine truth conditions for modal operators. AT is the set of atomic sentences.*

**(1)** $(\mathcal{M}_\lambda, \pi^i) \models p \in AT$ iff $p \in L(w_i)$.

**(2)** $(\mathcal{M}_\lambda, \pi^i) \models \neg A$ iff $(\mathcal{M}_\lambda, \pi^i) \not\models A$.

**(3)** Usual clauses for other Boolean connectives.

**(4)** $(\mathcal{M}_\lambda, \pi^t) \models B^t A$ iff $|A|^{\mathcal{M}_\lambda} \in E(w_t)$

**(5)** $(\mathcal{M}_\lambda, \pi^t) \models A > B$ iff $|B|^{\mathcal{M}_\lambda} \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$.

The only new element in this definition is the conditional neighborhood function C. If we select our attention to selections across co-present worlds, the idea when evaluating $A > B$ at $\pi^t$ is to focus on a class of alternative paths at $t$, $\pi'^t$ such that $A$ holds true in them. Then $\mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$ collects the events that hold true in all these alternative paths $\pi'^t$.

The selection process presupposed in the construction of the conditional neighborhood function $C$ can be seen, in similar terms than the ones suggested in [34]. First we can focus on selecting worlds with respect to the world $w_i = \pi^i$. Even if this selects the world $w_i$ uniquely, there is a second selection process that yields a class of paths $\pi'$ passing though $\pi^i$, and which necessarily include $\pi$ itself (validating $CMP$ immediately). For example, one can focus on the set of alternative paths that share with $\pi$ not only the same method of belief change but also other crucial aspects that methods of belief formation might share, like sameness in procedures for data-gathering, etc. Notice also that in this framework there is plenty of information external to the subject, which can also be used to identify the right set of alternative paths.[2]

So, the simplest way of constructing $\mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$ could proceed as follows: take the set of worlds $\pi'^t$ selected as we suggested above. Call this proposition $X$ and construct the conditional neighborhood as follows:

[2]Nothing precludes adopting here one of the most sophisticated solutions offered in [34] for defining conditional selection functions in the context of branching time, which abandons the conditional excluded middle – see page 316, [34]. This solution takes care of some of the many problems that arise when tense interacts with conditionals and is coherent with the application we are focusing on here.

$Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$ iff $X \subseteq Y$

This constructions would validate all the conditional properties we proposed above, and more. In fact, this construction will guarantee the *normality* of the generated conditional logic, and therefore it would satisfy as well some rules of inference, like *regularity* ($RCR$) which we did not impose above.

Semantically we have for every set of states $X, Y$ in $\mathcal{M}$ the following set of conditions for conditional neighborhoods:

**(cm)** If $X \cap Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$, then $X \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$, and $Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$.

**(cc)** If $X \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$, and $Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$, then $X \cap Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$.

**(cn)** $W \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$ .

The smallest classical, monotonic, and regular conditional logics are called $CE$, $CM$ and $CR$ respectively. The label $CK$ is reserved for the smallest normal conditional logic. $CM$ is determined by the class of conditional neighborhood models for which (cm) holds. $CR$ is determined by the calls in which both (cm) and (cc) hold and finally $CK$ is determined by the class of conditional neighborhood models satisfying all three conditions.

The syntactic conditions we used are weaker. For example consider the following option. Index every point $\pi^t$ with a finitely additive conditional probability function satisfying Dubins's axioms for primitive conditional probability [13].

Now we can determine that $Y \in \mathrm{C}(\pi^t, |A|^{\mathcal{M}_\lambda})$ if and only if $P_{\pi^t}(Y|\, |A|^{\mathcal{M}_\lambda})$ is greater than a threshold or directly one. It is clear that in this case (cc) is not satisfied and the that the resulting conditional system does not satisfy $RCR$. This kind of probabilistic model can be made compatible with other (doxastic) aspects of the general framework offered here by appealing to previous joint work [3]. In addition it satisfies all the conditional axioms we proposed (with the possible exception of $CMP$, which needs to be restricted in order to avoid nesting of conditionals in the consequent – and with an additional proviso that we will mention below).

In fact, the conditional probability function $P_{\pi^t}$ can be used to provide a more complete doxastic representation. Our doxastic neighborhoods should be augmented (given that our underlying notion of belief is an extension of $K$). One can then construct a *system of cores* in the neighborhood ordering points outside of the intersection of the neighborhood. We proposed in [3] seeing the outermost core in this construction as the *full beliefs* for $P_{\pi^t}$ and the innermost one as encoding the *almost certainties* or *expectations* for $P_{\pi^t}$. When the underlying probabilistic space is countable, the innermost core contains all the points of positive measure, and the cores order the points of zero measure.

This permits certain revisions with propositions of measure zero. As long as these propositions cut the outermost core, the revision (we show in [5]) is just the intersection of the proposition in question and the innermost core cutting it (if the space is infinite, but countable, and countable additivity is imposed then one can show that such innermost cutting core always exists, otherwise one needs alternative ways of proceeding). If the updating proposition is incompatible with the outermost core then the revision leads to incoherence or is undefined. The satisfaction of $CMP$ in this setting requires to impose an additional condition requiting that when the conditional probability of $B$ given $A$ is high or one, the material conditional $A \rightarrow B$ should be fully believed.

The extra degree of information given by having measures in doxastic neighborhoods could be useful in general, but we do not propose to use it to define conditionals probabilistically in order to investigate Nozick's notion of tracking truth. There are various reasons for this reluctance. In the first place it is known that the conditionals one obtains via this probabilistic analysis are both conceptually and formally rather different than the conditionals that Nozick uses in this characterization of knowledge. Textual evidence seems to indicate that syntactically the conditional system that attracted Nozick's attention is the one that David Lewis called **VW**, where the so-called 'centering' condition is replaced by 'weak centering'.

A system like this (or for that matter any other system of counterfactual conditionals in Lewis's hierarchy of conditional systems [26]) fails to satisfy the so-called export-import conditions. Nevertheless, in [4] is shown that the conditionals validated via probabilistic models of the sort sketched above do validate this law.

In addition, Nozick's definition might require engaging in couter-doxastic suppositions contradicting existing full beliefs. And these suppositions cannot be handled even by the most sophisticated notions of conditional probability admitting conditioning with events of measure zero. There are many other reasons supporting our reluctance. Nevertheless, formally our framework can make room both for probabilistic conditionals and for modeling non-normal conditionals probabilistic or not, which can be used in addition of the ontic conditionals that seem adequate to encode Nozick's ideas.

These ontic conditionals have little to do with the *epistemic* conditionals [6] that an agent might accept or reject at the epistemic states given by the epistemic function $E$. Assuming, for simplicity, the augmentation of the contents of the epistemic function, the evaluation of these conditionals requires supposing the antecedent of the conditional with respect to the intersection of the content of the epistemic function $E$, and then checking whether the consequent holds in this suppositional scenario. The recent discussion as to what are the adequate axioms for such a suppositional operation (see, the section on epistemic conditionals in [11] for an introduction to this problems and for detailed references) is relevant here in order to keep track of the (first person) reasoning capabilities of agents. Of course when such a stance is adopted the $\cap E$ has a completely different role: it is taken by the agent as the *standard for serious possibility* [24].

Treating an agent as a source of knowledge (i.e. in the sense in which knowing that P implies a license to testify as witness or authority [36], 489-523) leads to adopt an external stance with respect to it. In contrast we can be interested in representing the agent's view about his own knowledge. This leads to taking knowledge as evidence or a standard for serious possibility.

The logical underpinnings of both articulations of knowledge diverge as well. Here we have focused on the logical commitments of the external stance that Nozick himself adopted, where the focus of attention are knowledge attributions to an external agent rather than the knowledge claims of the agent itself. Nevertheless in a multi-agent environment the interplay between these two views of knowledge is also relevant, although its detailed consideration is still an open and relatively unexplored problem.

# References

[1] Arló-Costa, H. and Parikh, R. 'Some non-normal logics of knowledge without logical omniscience,' an abstract is forthcoming in the *Bulletin of Symbolic Logic*.

[2] Arló-Costa, H. and Pacuit, E. 'First Order Classical Modal Logic: Applications in logics of knowledge and probability,' TARK'05, (ed.) R. van der Meyden, NUS: Singapore, 262-79, 2005.

[3] Arló-Costa, H. and Parikh, R. 'Conditional Probability and Defeasible Inference,' *Journal of Philosophical Logic* **34**, 97-119, 2005.

[4] Arló-Costa, H. 'Bayesian Epistemology and Epistemic Conditionals: On the Status of the Export-Import Laws,' *Journal of Philosophy*, Vol. XCVIII, 11, 555-598, 2001.

[5] Arló-Costa, H. and Thomason, R. 'Iterative probability kinematics,' *Journal of Philosophical Logic* 46, 479-524, 2001.

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

[7] Bacharach, M.O.L., Gérard Varet, L.-A., Mongin, P. and Shin, H.S. (eds.) *Epistemic Logic and the Theory of Games and Decisions*, Theory and Decision Library, Vol. 20, Kluwer Academic Publishers, 1997.

[8] Barwise and L. Moss. *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena*, C S L I Publications, February 1996.

[9] Chellas, B. *Modal logic an introduction*, Cambridge University Press, 1980.

[10] Clarke, E. M., Emerson, E. A., and Sistla, A. P. 'Automatic verification of finite-state concurrent systems using temporal logic specifications,' *ACM Transactions on Programming Languages and Systems* 8 (2): 244-263, 1986.

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

[12] DeRose, K. 'Solving the skeptical problem,' *Philosophical Review*, 104, 1-52, 1995.

[13] Dubins, L.E. 'Finitely additive conditional probabilities, conglomerability, and disintegrations,' *Ann. Prob.* 3:89-99, 1975.

[14] Fagin, R., Halpern, J., Moses, Y., Vardi, M., *Reasoning about Knowledge*. The MITPress, Boston, 1995.

[15] Fine, K. 'Review of Lewis,' *Mind*, 84, 451-458, 1975.

[16] Halpern, J., van der Meyden, R., Vardi, M., 'Complete axiomatizations for reasoning about knowledge and time,' *SIAM Journal of Computing*, 33, 674 - 703, 2004.

[17] Halpern, J. 'Intransitivity and vagueness,' *Ninth International Conference on Principles of Knowledge Representation and Reasoning* (KR 2004), 121-129, 2004.

[18] Kelly, K. 'Learning Theory and Epistemology,' in *Handbook of Epistemology*, I. Niiniluoto, M. Sintonen, and J. Smolenski, eds. Dordrecht: Kluwer, 2004.

[19] Kracht, M and Wolter, F. 'Normal Monomodal Logics can simulate all others,' *Journal of Symbolic Logic*, 64(1), 99-138, 1999.

[20] H. E. Jr. Kyburg. 'Probabilistic inference and non-monotonic inference,' *Uncertainty in Artificial Intelligence*, R.D Schachter, T.S. Evitt, L.N. Kanal J.F. Lemmer (eds.), Elsevier Science (North Holland), 1990.

[21] Kyburg, H. E. Jr. and Teng, C. M. 'The Logic of Risky Knowledge,' *Proceedings of WoLLIC*, Brazil, 2002.

[22] Kupferman, O. and Grumberg, O. 'Branching-time temporal logic and tree automata,' *Information and Computation*, 125(1):62-69, 25 February 1996.

[23] Lehmann, D., Magidor, M.: 1992, 'What does a conditional base entails?' Artificial Intelligence, 55, 1-60.

[24] Levi, I. *The Enterprise of Knowledge*, MIT Press, Cambridge, Mass., 1980.

[25] Lewis, D. 'Elusive Knowledge,' *Australasian Journal of Philosophy*, 74, 4, 1996.

[26] Lewis, D. *Counterfactuals*, Harvard University Press, Cambridge, Mass., 1973.

[27] Montague, R. *Universal Grammar*, *Theoria 36*, 373–98, 1970.

[28] Nozick, R *Philosophical Explanations*, Oxford University Press, 1981.

[29] Parikh, R. 'Logical omniscience and common knowledge: WHAT do we know and what do WE know?, TARK'05, op cit., 62-78.

[30] Parikh, R., and Ramanujam, R. 'A Knowledge based Semantics of Messages,' *J. Logic, Language and Information* **12** 2003, 453-467.

[31] "A Counterexample to Modus Ponens," *Journal of Philosophy*, LXXXII, 9 (Sep., 1985).

[32] Roush, S. *Tracking Truth: Knowledge, Evidence and Science*, Oxford, 2005.

[33] Scott, D. 'Advice in modal logic,' K. Lambert (Ed.) *Philosophical Problems in Logic*, Dordrecht, Netherlands: Reidel, 143–73, 1970.

[34] Thomason, R. and Gupta, A. 'A theory of conditionals in the context of branching time,' in *Ifs*, W.L. Harper, R. Stalnaker and G. Pearce (eds.), Reidel, 299-322, 1980.

[35] van Benthem, J. 'Epistemic logic and epistemology: the state of their affaires,' *Philosophical Studies*, forthcoming.

[36] Williamson, T. *Knowledge and its limits*, Oxford, 2000.