CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Almae Matris Studiorum Campus

G. Corsi
E. Orlandelli

Free Quantified Epistemic Logics

Final draft. Published version in *Studia Logica 101: 1158–1183 (2013)*
The final publication is available at link.springer.com,
http://link.springer.com/article/10.1007%2Fs11225-013-9528-x#page-1

Abstract. The paper presents an epistemic logic with quantification over agents of knowledge and with a syntactical distinction between *de re* and *de dicto* occurrences of terms. Knowledge *de dicto* is characterized as 'knowledge that', and knowlegde *de re* as 'knowledge of'. Transition semantics turns out to be an adequate tool to account for the distinctions introduced.

Keywords: first-order epistemic logic, multi-agent systems, term-modal language, indexed modalities, transition semantics.

1. Introduction

Propositional epistemic logics are extensively used for representing agents' knowledge,[1] the basic idea being to express the knowledge of agents $1, \ldots, n$ by means of a finite set of modalities, $\mathcal{K}_1, \ldots, \mathcal{K}_n$, to be interpreted in a possible-world semantics with the help of n accessibility relations. One of the advantages of modeling knowledge by means of propositional epistemic logic consists in the modularity of such approach inasmuch as one and the same semantic environment can serve to model different notions of knowledge or belief just by varying the algebraic structure of the accessibility relation(s). So, after having established the algebraic structure appropriate for the epistemic notion (knowledge, belief, etc.) that one wants to model, an agent's knowledge about the world and about other agents' knowledge can be expressed by formulas such as $\mathcal{K}_i A$ and $\mathcal{K}_i \mathcal{K}_j B$, respectively.

Although such an approach can be adequate in a number of cases, it encounters serious limitations because it doesn't allow us to model agents' knowledge about possibly infinite sets of objects and their properties and for various applications the expressive power of quantified epistemic logics is necessary. In his book *Knowledge and Belief*, 1962, Hintikka discusses at length the interplay between epistemic modalities and quantifiers. An immediate gain obtained by merging the language of classical quantified logic with that of an n-modal propositional epistemic logic is that we are able to express an agent's knowledge about particular objects, *agent 1 knows that s is brave*, $\mathcal{K}_1 \mathsf{Brave}(s)$, as well as about general objects, *agent 1 knows that someone is brave*, $\mathcal{K}_1 \exists x \mathsf{Brave}(x)$, or *Someone is known by agent 1 to be brave*, $\exists x \mathcal{K}_1 \mathsf{Brave}(x)$.

[1] See e.g. [6], [18] and [10].

Still a fundamental limitation remains since we are dealing with a fixed set of agents. Now, in order to represent a system in evolution, e.g. a network where out-of-date elements (agents) are substituted by new ones, we need to reason about the epistemic states of a dynamic set of agents. In general, we want to deal with a variable, indefinite, and possibly infinite set of agents and to reason about them from inside the system itself. So it is natural to assume that the agents are members of the domain of objects we are talking about, have names[2] inside the system and are quantified over. All this is in tune with: "To make epistemic logic pertinent to epistemology, computer science, artificial intelligence and cognitive psychology the agent must be activated. The original symbolic notation of a knowing agent also suggests this: an agent should be inside the scope of the knowledge operator - not outside, as Hintikka notes in [15]" ([13]: 7).

In conclusion, we may need to refer to generic agents, 'someone knows that s is brave', as well as to groups of agents, 'every friend of s knows that s is brave'. Finally we want agents to reason about themselves as in 'everyone knows if he or she is in pain'. Almost all the situations described so far were noted as early as 1951 when, in the final note to the section on epistemic logic of *An Essay in Modal Logic*, G.H. von Wright claimed that

> [w]e could develop an alternative system in which the epistemic modalities are treated as 'relative' to persons. In this system we should have to deal with expressions like 'known to x', 'unknown to x', *etc.* Introducing quantifiers we should get a combined system dealing with expressions like 'known to somebody', 'unknown to everybody', *etc.* This combination of epistemic and existential modalities will not be studied in the present essay. ([20]: 35.)

Von Wright's insight has, to our knowledge, remained unnoticed until recent times when term-modal logics have been proposed in [11], [3], and [9].[3] In the following with the expression 'term-modal logics' we denote not only the logics described in the paper *Term-Modal Logic*, [9], but also the logics discussed in [3] and [11]. The reason for this is that, despite the quite different points of view, all of them are characterized by the fact that the epistemic operators are indexed by terms of the language and quantifying over the agents is allowed. In this new context we can express

- knowledge of a particular agent t, *t knows that A*, $|t|A$
- knowledge of a generic agent, *someone knows that A*, $\exists x|x|A$

[2]An implicit assumption usually made is a one-to-one correspondence between agents and agents' names: given that an agent is nothing more than an index to which an accessibility relation is associated, we must assume that different indices (i.e. agents names) refer to different agents (i.e. indices), and that one and the same index always refers to the same agent. This implies, at least from a meta-theoretical point of view, that a formula like $\mathcal{K}_a\mathcal{K}_b P \wedge \mathcal{K}_a\mathcal{K}_c P$ implies that *a knows of two* different *agents that they know that P*.

[3]In [14], although this possibility is mentioned in at least two passages, the formation rules for formulas don't allow for it.

- knowledge of groups of agents,
  - *some policeman knows that* A, $\exists x(P(x) \wedge |x|A)$
  - *every policeman knows that* A, $\forall x(P(x) \rightarrow |x|A)$

2. Towards a more general approach

Although term-modal languages constitute a major step in the development of quantified epistemic logics we believe that they are a starting point rather then the end of the story. A crucial peculiarity of the term-modal operator $|\,t\,|$ is that it is invariant with respect to the formula that follows it: $|\,t\,|Pc$ is well formed as well as $|\,t\,|Px$, see [9]. An immediate consequence is that it is impossible to distinguish between 't *knows that* Pc' and 't *knows of* c *that it is* Px'. A typical way out in this situation is to resort to the λ-abstraction operator so as to distinguish between: $|\,t\,|(\lambda x(Px).c)$ and $\lambda x(|\,t\,|(Px)).c$, respectively.[4] We aim at a language able to handle not only problems of scope, but also the interplay between quantifiers and modalities and for this further issue the presence of the λ-operator is of no help.[5] Let us start by reviewing some of the questions much discussed in the literature.

2.1. *'Knowing of'* and *'knowing that'*.

As is well known from [19], the sentence:

(1) Ralph knows that someone is a spy

may be read either as 'Ralph knows that there are spies' or as 'Ralph knows of someone that he is a spy'.[6] As Quine says, "the difference is vast; indeed, if Ralph is like most of us, ['Ralph knows that there are spies'] is true and ['Ralph knows of someone that he is a spy'] is false" ([19]: 178).[7] Furthermore, also epistemic sentences apparently not involving quantifiers seem to be ambiguous between two alternative readings, witness:

(2) Ralph knows that Ortcutt is a spy

which may be rephrased either as 'Ralph knows that Ortcutt, whoever he is, is a spy' or as 'Ralph knows of Ortcutt that he is a spy'.[8]

The distinction between the two readings of sentences such as (1) and (2) is usually referred to in modal contexts as a distinction between two kinds of

---

[4] See, e.g., [7] or [8].

[5] See section 5.

[6] In [19] this reading is rendered by 'there is someone whom ralph knows to be a spy'. We take this rendering, as well as the passive one 'someone is known by Ralph to be a spy', as equivalent to that in terms of 'knowing of'.

[7] Quine's original example is an attribution of belief, not of knowledge. But this difference is inessential for the present purposes.

[8] Where 'he' is not a pronoun of laziness, see [1].

modality, *de dicto* and *de re*. Roughly, in the epistemic contexts, we may want to distinguish between expressing a relation between an agent and a sentence,[9] and expressing a relation among an agent, an object, and an open formula, respectively. This is not the place for a thorough philosophical discussion of the nature of these two kinds of knowledge; suffice it to say that an epistemic sentence seems to have at least two readings, *de dicto* and *de re*, that neither of the two seems to involve quantifiers essentially, and that the syntactic difference seems to boil down to that between formulas involving operators occurring before closed formulas and operators occurring before open ones. The term-modal language of [9] is unable to express this distinction which is basically a distinction between 'knowing something *of* some object' and 'knowing *that* such and such a sentence (is true)'.[10]

To reach this goal, we introduce complex term-modal operators of the form $|t : {}^{c}_{x}|$ that allow for formulas such as:

$$|t : {}^{c}_{x}|P(x)$$

t knows of c that (s)he is P(x)

Here c is any term and it takes outer scope, so that that formula expresses of the actual denotation of c that it is known by t to satisfy $P(x)$. This formula should be contrasted with:

$$|t : \star|P(c)$$

which expresses the fact that t *knows that* $P(c)$. $\star$ stands for the empty sequence of variables.[11] In order to grasp the notation above, think of $|t : {}^{c}_{x}|P(x)$ and $|t : \star|P(c)$ as analogous to $\lambda x(|\,t\,|(Px)).c$ and $|\,t\,|(\lambda x(Px).c)$, respectively.

Knowing of and *agent-denoting terms.* The issue of distinguishing *knowing of* from *knowing that* becomes even more pressing when we have to account for nested operators with agent-denoting terms. It may make a big difference whether *I know* of *t that he knows* ... or whether *I know* that *t knows* .... To illustrate, we ask the reader to consider the following scenario proposed by Grove:

> Suppose there are two robotic agents, A and B, and A has just broken down. He sends a cry for help over a public broadcast system. B, who is the agent responsible for dealing with such matters, may or may not have heard. So A's subsequent action depends on whether

[9] Or its semantic content. We don't want to take a stance on this question.

[10] The distinction we are proposing should not be conflated with that proposed in [14] between 'knowing who' and 'knowing that'. We believe that so-called wh-questions may fall under any one of the two notions of knowledge we are dealing with. Suppose, for example, that I'm asked who is Obama. While in some contexts, say at an exam at school, in order to answer it I have to know *that Obama is the president of the US*, in some other context, say at a party at the White House, what is needed is knowledge *of someone in particular that he is Obama*. It is quite clear that in the two scenarios I'm asked for different kinds of knowledge, but also that each is an example of wh-knowledge.

[11] Henceforth we will use 'know that' to express the *de dicto* reading.

> he can deduce that "I (A) know that B knows that I need help" (If this is true he can just wait, but otherwise he should try something else). So what is a good formalization of "I (A) know that B knows that I need help"? ([11]: 314)

For a start let us observe that our formalism allows various renderings of the sentence in question:

(3) outer scope for both a and b

$$|a : \substack{a\\y}\ \substack{b\\x}||x : y|\mathsf{Needhelp}(y)$$

a knows of himself and of b that the latter knows of the former that he needs help.

(4) outer scope for a and not for b

$$|a : \substack{a\\y}\ ||b\ : y|\mathsf{Needhelp}(y)$$

a knows of himself that b knows of a that he needs help.

(5) outer scope for b and not for a

$$|a : \substack{b\\x}\ ||x\ : \star|\mathsf{Needhelp}(a)$$

a knows of b that he knows that a needs help.

(6) outer scope for neither a nor b

$$|a : \star||b : \star|\mathsf{Needhelp}(a)$$

a knows that b knows that a needs help.

The answer to Grove's question depends on the specific epistemic context one wants to describe. The crucial point is to have a semantical framework in which the renderings (3)-(6) may have distinct semantic values.

Syntactic characterization of de re. We have introduced an overtly syntactic distinction between *de dicto* and *de re* epistemic operators. Such a distinction doesn't amount to a distinction of scope between quantifiers and epistemic operators as in the Russellian analysis. One advantage of our rendering is that it allows to express both readings without having to introduce some quantifier where there seem to be none - *e.g.* for the *de re* reading of (2).[12] Still one may wonder whether there is some independent motivation for our approach. We believe that it may be seen as parallel to Quine's multigrade treatment of verbs of propositional attitudes. The general form of our operators - *i.e.* $|t : \substack{t_1\\x_1} \ldots \substack{t_n\\x_n}|$

[12]Note that this is achievable also by introducing the operator λ.

- is analogous to Quine's predicate $Bel(t, \vec{t}, A(\vec{x}))$, where t stands for the subject of the attitude, $\vec{t}$ for the n-ary vector of objects about which the attitude is, and $A(\vec{x})$ for an n-ary propositional function/open formula. We agree that "[t]he key to the *de re/de dicto* distinction, as we are representing it, is explicit in these formulation. $[|t : \star|]$ applies to what is expressed by a closed sentence; $[|t : \frac{\vec{t}}{\vec{x}}|]$ applies in part to what is expressed by an open sentence and in part to a *res*" ([1]: 343).

2.2. *De dicto* and *de re* knowledge

Some semantic desiderata. We have thus far enriched the term-modal operators to distinguish between *knowing that* and *knowing of.* Now we have to supplement our syntactic distinction with a semantics that captures it.[13] One basic desiderata is to have a semantics where a formula such as $|t : \frac{\vec{s}}{\vec{x}}|A(\vec{x})$ expresses a relation between the agent denoted by t, the objects denoted by members of $\vec{s}$, and the formula $A(\vec{x})$. Given that we want the *de dicto* and the *de re* readings of a sentence such as (2) to differ, the semantic value of $|r : \star|S(o)$ should differ from that of $|r : \frac{o}{x}|S(x)$. $|r : \star|S(o)$ should express a relation between an agent and a sentence, whereas $|r : \frac{o}{x}|S(x)$ between an agent, an object, and an open formula. Furthermore, if we want to retain Quine's distinction between opaque and transparent occurrences of a term in an epistemic formula, only the occurrence of o in $|r : \frac{o}{x}|S(x)$ is substitutable, but not that in $|r : \star|S(o)$.

Compatibility relation. If possible world semantics is to be capable of modeling epistemic attitudes, it is natural to think of possible worlds as states of affairs compatible with the epistemic state of some agent and guided by this idea we are led to replace the usual accessibility relation between worlds by a compatibility relation between agents and worlds. To this aim we introduce a binary relation of *compatibility*, $a \prec v$, meaning that world v is compatible with the epistemic state of agent a. Similarly in [9], a ternary relation is introduced between a world w_1, an agent of that world and another world w_2, $w_1 \xrightarrow{a} w_2$, meaning that w_2 is *a-reachable* from w_1. We refrain from explicitly indicating which world a inhabits without fear of confusion, as we shall see.

Rigidity of terms. As is well known, in possible world semantics a term is said to be *rigid* if and only if it denotes the same object in every accessible/compatible world. With rigid terms we cannot express any interesting notion of knowledge related to identities, in fact every agent knows every true identity statement, thus depriving identity statements of their relevance. Just to cite a well known example, if h stands for Hesperus and p stands for Phosphorus at w, we can have that $h = p$ is satisfied at w. Why should this entail that each agent knows, at w, *that* Hesperus and Phosphorus are identical? To avoid this undesirable consequence, we need a semantics with non-rigid designators, that is a semantics where two individual constants, say h and p, that denote the same object in w can denote two different objects in some world w_1 that is

[13] In taking the possiblity of *de re* knowledge at face value, we diverge from the term-modal logic presented in [11] where all knowledge is *de dicto*.

compatible with the epistemic state of some agent a in w. In this way an agent may not know *that* Hesperus and Phosophorus are identical.

The semantics of de dicto knowledge. Now that we have introduced the compatibility relation, and the need for non-rigid designators, we are able to give our semantics for a *de dicto* formula such as $|r:\star|S(o)$: this formula is true at a world w if and only if the embedded sentence $S(o)$ is true at every world w_i that is compatible with the epistemic state of the agent denoted by r in w. So, for example, if the formula $|r:\star|S(o)$ represents 'Ralph knows that Ortcutt, whoever he is, is a spy', our semantics says that that sentence is true if and only if at every Ralph-compatible world, the denotation of 'Ortcutt' in that world is a spy. Note that the denotation of o has to be determined locally in the worlds w_i that are compatible, and not once and for all in w. If names were to be interpreted as rigid designators, this qualification would have been irrelevant, but this interpretation is unreasonable in an epistemic context particularly when confronted with nested operators, as already mentioned. Take 'x believes that C.L.Dodgson believes that Alice is nice', the denotation of 'C.L.Dodgson' might depend on how x represents to himself C.L.Dodgson: is he a writer of children's stories or is he a mathematician?[14]

Transition relation. By having non-rigid designators we have ensured that the identity of Hesperus and Phosphorus doesn't entail that each agent knows *that* Hesperus and Phosphorus are identical, but what in the case of *de re* knowledge? Does the identity of Hesperus and Phosphorus entails that each agent knows *of* Hesperus and Phosphorus that they are identical? If we were to formalize the distinction between the two readings of a sentence such as (2) simply as being a consequence of nonrigid terms,[15] as is customary in possible-world semantics, then the answer to this question would be yes. But we believe that it should be no.

In fact we depart from possible world semantics and arrive at transition semantics. Generally speaking, the main feature of transition semantics (or counterpart semantics) consists in the existence of a relation between elements of different domains (objects inhabiting different worlds) so that in order to establish if 'c *is necessarily* A' one has to take into account all the transitions/counterparts of c in all accessible worlds. When we deal with alethic modalities, the counterpart relation is absolute, given from the outset, see [4]. Now transition semantics seems to adapt very well to epistemic contexts as long as the transition relations are made agent-dependent. To this end we introduce for each agent a a transition relation $b \overset{a}{\rightarrowtail} c$ between objects inhabiting different worlds, meaning that c is an a-transition of b. We think of the transition relation $b \overset{a}{\rightarrowtail} c$ as a relation between a *representative* and a *representee*: for all a knows of b, c may be/represent b in the a-compatible world that he inhabits. And, led by this idea, we propose the following semantics for a *de re* formula such as $|r:{}^{o}_{x}|S(x)$: this formula is true at a world w if and only if at every

[14]Example taken from [3].

[15]So that in a *de re* formula some term has wider scope than the operator, whereas in a *de dicto* one it's the operator that has wider scope.

r-compatible world, the open formula $A(x)$ is true of every local representative of b —*i.e.* of every c such that $b \overset{r}{\rightarrowtail} c$.

Acquaintance. According to our proposal the notion of *de re* knowledge should express some relation between an agent, some object, and some open formula. One of the main tenets of the Quine-Kaplan-Lewis analysis of *de re* knowledge[16] is that an agent's *de re* knowledge of an object depends on there being a relation of *acquaintance* between the agent and the object, where the relation of acquaintance is some epistemological relation occurring between a and b that enables a's piece of knowledge to be *about/caused by* b. If we think of the set of all c such that $b \overset{a}{\rightarrowtail} c$ as being dependent on the relation of acquaintance, then we are proposing a semantics for *de re* knowledge that follows to some extent the Quine-Kaplan-Lewis analysis: whether a knows something *of* b depends on an acquaintance relation between a and b.[17]

Agent-dependency. One central feature of the epistemic transition relation is that it is agent-dependent: suppose that w_1 is compatible with the epistemic states of two agents both inhabiting w; whether some c inhabiting w_1 is a representative of some object inhabiting w is a question that depends on which of the two agents we are considering. Consider the case of the twins a_1 and a_2, both inhabiting w, that are equivalent w.r.t. *de dicto* knowledge: every a_1-compatible world is also a_2-compatible, and vice versa. If their respective *de re* knowledge depends on some agent-independent relation, such as trans-world identity, then we must conclude that they have the same *de re* knowledge. The problem is that if *de re* knowledge depends solely on the compatibility relation, a_1 knows something of some object if and only if a_2 knows it. So they can't be differentiated by their respective *de re* knowledge.[18] With the help of transitions we block this unwanted conclusion because *de re* knowledge depends both on the compatibility and on the transition relations, but nothing in our example entails that $b \overset{a_1}{\rightarrowtail} c$ *iff* $b \overset{a_2}{\rightarrowtail} c$.

Transitions and identity. The use of transitions, instead of trans-world identities, allows also to deal successfully with the following scenario presented by Carlson:

> Let's consider
>
> (29) There is someone who might be two different people as far as the police know.

[16] The label is borrowed from [12].

[17] Where we diverge, especially from Kaplan's proposal in [16], is in neglecting that a may know *of* b both $A(x)$ and $\neg A(x)$. In our terminology Kaplan takes the relation of aquaintance as existential quantification over representatives, whereas we take it as universal quantification. One way of reducing the divergence is by claiming that we have "replace[d] existential quantification over aquaintance relations by reference to a contextually salient particular acquaintance relation" ([12]: 210).

[18] This example is somehow analogous to Lewis' 'two gods case', see [17]. The main difference is that Lewis point was meant to distinguish *de dicto* from *de se*, whereas we are distinguishing *de dicto* from *de re*.

> The someone in (29) could be a person who leads a double life, like Arthur Brownjohn alias Major Elsonby Mellon in Julian Symon's novel *The Man who killed himself.* Arthur Brownjohn deceives the police about his double identity so successfully that he ends up accused of the murder of his alter ego. Here, apparently, an individual who is being traced in an the actual situation according to the epistemic alternatives of the police does in fact split into two imaginary individuals: what the police know does not rule out the possibility that Arthur Brownjohn should have killed Major Elsonby Mellon and walked away from it, in which case the two individuals cannot very well be one and the same. On the other hand, the identification of each of them separately with Arthur is justified by the fact that the police have been intentionally misled to construe one real world person as two people. ([2]: 238)

(29) can be formalized as

$$\exists x \exists y (x = y \wedge \forall z (Policeman(z) \rightarrow \langle z : xy \rangle x \neq y))$$

Given that the story presented by Carlson appears to be a real possibility, such a formula should be satisfiable. In transition semantics this is accomplished by allowing an object to have none, one or many transitions w.r.t. an agent in one and the same compatible world. Thus even if the objects denoted by x and y coincide in the actual world, they can diverge in some world that is compatible with what the police know in the actual world.

De re vs *de dicto again.* Suppose we say[19] that a world v is compatible with the epistemic state of agent a (inhabiting w) if all the sentences that a believes are true in v (in other terms if the *de dicto* knowledge of a is verified in v), then it may well be that the same worlds are compatible with the epistemic state of two different agents a and b: they share the same *de dicto* knowledge. Still their *de re* knowledge may differ because the truth conditions of *de re* sentences make use of the transition relations which are agent-dependent. This is a point that needs some clarification. In general, one of the ways in which agents relate themselves to the world is by identifying objects, and often they make mistake in doing so. Transitions can be seen as ways of identification. A possible world compatible with an agent a is just a way in which the world could be as far as a knows. If in the actual world the agent a identifies Venus in two different ways, then in worlds compatible with what a knows of the actual world there will be two different objects that represent the two different identifications of Venus. Notice that agents associate objects to other objects without any recourse either to linguistically triggered ways of presentation or to the names of the objects in question. Knowledge *de re* is knowledge under every representative of the object about which the attitude is. In conclusion, an agent a knows an (open) formula $A(x)$ *of* some b just in case in every a-compatible world, every a-transition of b - *every way that b may be that is compatible with a's knowledge* - satisfies

[19] As in fact we say when building canonical models, see [5].

$A(x)$. As to *de dicto* knowledge, our approach is parallel to that of possible-world semantics with non-rigid terms: to evaluate an epistemic sentence, we must first move to a-compatible worlds, and then determine the denotation of the terms.

2.3. *Agents, domains, worlds, and rigidity.*

Agents and domains. Given that the compatibility relation is between agents and worlds, we should be able to partition the objects of a domain between those which are agents and those which are not. A way out would be to have two sorts of objects: agent and non-agents. The fact is that by doing so we encounter a plethora of problems which are more pertinent to the possible particular applications than to the general framework we are trying to build. In [3] this distinction is maintained, and it leads to a truth-value-gap semantics. In the present context we could operate this distinction by saying that an element d of a domain is said to be an agent if there is a v such that relation $d \prec v$ holds. Thus a non-agent is an object whose epistemic state is not compatible with any world and therefore is an object that believes everything, also contradictions. In our language however we do not introduce any partition among the individual constants which refer to agents and those which don't, so any term can take the place of an agent inside an epistemic operator. This has the advantage that no special restrictions on the formation of formulas are needed.

Worlds and Domains. For the sake of generality we make a further assumption, that any world w is endowed with two domains, an inner domain D_w and an outer domain U_w such that $D_w \subseteq U_w$ and $U_w \neq \emptyset$. The reason for this choice is basically due to an analogy with quantified modal logics with actualist quantification. It is not immediately obvious what the traditional distinction between existing (D_w) and possible (U_w) objects could corresponds in an epistemic context, nevertheless we want to keep the domain of variation of the quantifiers, D_w, separate from the domain of interpretation of the variables and of the descriptive symbols of the language, U_w. This allows us to admit of empty domains (D_w may be empty) and so to get quantified epistemic logics based on free logic instead of classical logic. Furthermore, we try to build up a semantics that doesn't imply the validity of any controversial modal principle and the double domain enables us to falsify the so-called converse of the Barcan formula: I may believe that 'all the basket players are taller than myself' simply because I am not aware of all of them, for example I am not aware of Muggsy Bogues, so why should I be compelled by the logic to conclude that 'I believe of each basket player that he is taller than myself'? As usual in quantified modal logics, a choice has to be made as to the inner domains: should they be constant, increasing, decreasing, varying? We opt for varying domains and we leave to the different applications to determine the appropriate constraints. For the sake of simplicity we assume that domains are pairwise disjoint. In any case this is not a real limitation because we can always label an object by the world it inhabits.

a-rigidity and a-stability. In our semantics every term denotes in every state

of affairs (world), and it can denote different objects in different worlds. Let us agree that the name c denotes d_1 in w_1 and d_2 in w_2. For simplicity's sake let us assume that w_2 is the only world compatible to w_1 according to the epistemic state of an agent a which inhabits w_1. Every element d_1 of the domain of w_1 can have, in general, multiple a-transitions in the domain of w_2. What can a-rigidity mean in this context? We borrow the analogous notion from modal logics, see [4], and we say that c is *a-rigid* if the interpretation of c in w_2, i.e. d_2, is one of the a-transitions of d_1 in w_2. We are quite far from the traditional notion of rigidity, according to which c is rigid only if $d_1 = d_2$. We can try to approximate it by imposing that the set of a-transitions of d_1 is either the empty set or a singleton. Under this proviso, an a-transition of d_1 coincides with the interpretation of c in w_2. When this happens we say that c is *a-stable*. We can envisage a situation in which certain names are rigid with respect to certain agents and not rigid with respect to others. What it is worth noticing at this point is that the notions of a-rigidity and of a-stability correspond to the validity of particular formulas, see section 4.0.1, so we can have formal systems that include at the same time rigid and non-rigid terms.

3. Language

Consider a first-order signature with identity, individual constants as well as predicate and function symbols. Terms are defined as usual. The primitive logical symbols are $\bot, \rightarrow, \forall$, and, for $n \geq 0$, $|t : \frac{t_1}{x_1} \dots \frac{t_n}{x_n}|$, where $x_1, \dots, x_n$ are pairwise distinct variables and $t, t_1, \dots, t_n$ are terms. When $n = 0$, we write $|t : \star|$. The other connectives and the existential quantifier are defined as usual. $|t : x_1 \dots x_n|$ stands for $|t : \frac{x_1}{x_1} \dots \frac{x_n}{x_n}|$, and $< t : \frac{t_1}{x_1} \dots \frac{t_n}{x_n} >$ for $\neg|t : \frac{t_1}{x_1} \dots \frac{t_n}{x_n}|\neg$. We will also use $\vec{x}$ and $\vec{t}$ for arrays of variables and terms, respectively.

The definitions of formula and of free variable are the usual ones for first-order language augmented by the following:

- If $A(x_1, \dots, x_n)$ is a formula whose free variable are among $x_1, \dots, x_n$, then $|t : \frac{t_1}{x_1} \dots \frac{t_n}{x_n}|A$ is a formula, where $t, t_1, \dots, t_n$ are terms. The free variables of $|t : \frac{t_1}{x_1} \dots \frac{t_n}{x_n}|A$ are all (and only) the variables occurring in $t, t_1, \dots, t_n$.

The proviso that all the free variables occurring in the scope of an epistemic operator must occur inside the operator implies that the free variables always take outer scope.

Substitution The approach we propose inherits many features from the counterpart semantics for quantified modal logics introduced in [4]. In particular it inherits the point of view that the focus of many problems and of their solutions resides in the principle of substitution. This principle in term-modal logic is formulated as:

$$(|t|A)[s/y] \leftrightarrow |t|(A[s/y])$$

It can be seen as saying that the scope is immaterial: the term s in $(|t|A)[s/y]$ has outer scope and in $|t|(A[s/y])$ has inner scope, still the two formulas are equivalent. We give a definition of substitution that differs from the standard one just for the modal case for which it holds that:[20]

$$(|t : {}^{t_1}_{x_1} \cdots {}^{t_k}_{x_k}|B)[\vec{s}/\vec{y}] =_{df} |t[\vec{s}/\vec{y}] : {}^{t_1[\vec{s}/\vec{y}]}_{x_1} \ldots {}^{t_k[\vec{s}/\vec{y}]}_{x_k}|B$$

where $\vec{y}$ contains all the free variables occurring in $t, t_1, \ldots, t_k$.
Notice that substitutions are performed inside the modal operator and not inside the formula that follows it. This makes the whole difference: we do not allow substitutions to commute with epistemic operators by virtue of the definition of substitution. Instances of the commutativity of substitutions correspond to precise semantical conditions that can be assumed or not assumed as the case may be.

4. Semantics

Definition 4.1. *An* epistemic transition structure[21] *is a quintuple* $\mathcal{F} = \langle \mathcal{W}, \mathcal{U}, \mathcal{D}, \prec, \rightarrowtail \rangle$*, where:*

- $\mathcal{W} \neq \emptyset$ *is the set of possible worlds.*
- $\mathcal{U} = \{U_w : w \in \mathcal{W}\}$ *is a family of non-empty [and pairwise disjoint] sets, we will refer to* U_w *as the* outer domain *of* w*. Each* U_w *is the (world-relative) domain of interpretation of the signature of* $\mathcal{L}$*. If* $a \in U_w$*, we say that* a inhabits w*.*
- $\mathcal{D} = \{D_w : w \in \mathcal{W}\}$ *is a family of (possibly empty) sets such that* $D_w \subseteq U_w$*, we will refer to* D_w *as the* inner domain *of* w*. Each* D_w *is the (world-relative) domain of quantification. If* $a \in D_w$*, we say that* a exists *in* w*.*
- $\prec \subseteq (\mathcal{U} \times \mathcal{W})$ *is the* compatibility relation *between agents and worlds. A world* v *is* a-compatible *whenever* $\langle a, v \rangle \in \prec$*.*
- $\rightarrowtail = \bigcup_{a \in \mathcal{U}} \{\overset{a}{\rightarrowtail}\}$*, where for all* $a \in \mathcal{U}$*,* $\overset{a}{\rightarrowtail} \subseteq \{a \times U_w \times U_v : a \in U_w \,\&\, a \prec v\}$ *is the agent-dependent transition relation. If* $\langle a, b, c \rangle \in \overset{a}{\rightarrowtail}$*,* c *is said to be an* a-transition *of* b *and we usually write* $b \overset{a}{\rightarrowtail} c$*.*

Definition 4.2. *Let* $\mathcal{F} = \langle \mathcal{W}, \mathcal{U}, \mathcal{D}, \prec, \rightarrowtail \rangle$*. An* epistemic transition model *based on* $\mathcal{F}$ *is a pair* $\mathcal{M} = \langle \mathcal{F}, \mathcal{I} \rangle$*, where* $\mathcal{I}$ *is a function associating to any* $w \in \mathcal{W}$ *an interpretation* I_w *such that:*

- *for any constant* c*,* $I_w(c) \in U_w$*,*

[20] The full definition is spelled out in [5].

[21] We generalize the semantics introduced in [5] in that we introduce double-domains. In [4] transition semantics was given for a mono-modal language with indexed operators.

- *for any function* f^n, $I_w(f^n) : (U_w)^n \to U_w$,
- *for any relation* P^n, $I_w(P^n) \subseteq (U_w)^n$,
- $I_w(=) =_{df} \{\langle a, a, \rangle : a \in U_w\}$.

DEFINITION 4.3. *For any* $w \in \mathcal{W}$*, a* w-assignment *is a world-relative function* $\sigma : Var \to U_w$*. Given a* w*-assignement* σ*, by* $\sigma^{x \triangleright d}$ *we denote the* w*-assignment that behaves exactly as* σ *save for the variable* x *that is mapped to* $d \in U_w$*.*

DEFINITION 4.4. *Given an interpretation* $\mathcal{I}$ *of* $\mathcal{L}$ *over* $\mathcal{F} = \langle \mathcal{W}, \mathcal{U}, \mathcal{D}, \prec, \rightarrowtail \rangle$ *and a* w*-assignment* σ*, the interpretation of* t *in* w *under* σ*,* $I_w^\sigma(t)$*, is defined as:*

- $I_w^\sigma(x) =_{df} \sigma(x)$
- $I_w^\sigma(c) =_{df} I_w(c)$
- $I_w^\sigma(f(t_1, \ldots, t_n) =_{df} I_w(f)(I_w^\sigma(t_1), \ldots, I_w^\sigma(t_n))$

When no ambiguity arises, we will write $\sigma(t)$ *as a shorthand for* $I_w^\sigma(t)$*.*

DEFINITION 4.5. *We inductively define satisfaction of a formula* A *at* w *under* σ *in* $\mathcal{M}$*,* $\sigma \models_w^{\mathcal{M}} A$*, as follows:*

$$\begin{array}{lcl}
\sigma \not\models_w^{\mathcal{M}} \bot & & \\
\sigma \models_w^{\mathcal{M}} s = t & \Longleftrightarrow & \sigma(s) = \sigma(t) \\
\sigma \models_w^{\mathcal{M}} P^k(t_1, ..., t_k) & \Longleftrightarrow & \langle \sigma(t_1), ..., \sigma(t_k) \rangle \in I_w(P^k) \\
\sigma \models_w^{\mathcal{M}} B \to C & \Longleftrightarrow & \sigma \not\models_w^{\mathcal{M}} B \textit{ or } \sigma \models_w^{\mathcal{M}} C \\
\sigma \models_w^{\mathcal{M}} \forall x B & \Longleftrightarrow & \textit{for every } d \in D_w,\ \sigma^{x \triangleright d} \models_w^{\mathcal{M}} B \\
\sigma \models_w^{\mathcal{M}} |t : {}^{t_1}_{x_1} \cdots {}^{t_n}_{x_n}| B & \Longleftrightarrow & \textit{for all } v \textit{ such that } \sigma(t) \prec v, \\
 & & \textit{for all } v\textit{-assignment } \tau, \textit{ if} \\
 & & \sigma(t_1) \overset{\sigma(t)}{\rightarrowtail} \tau(x_1) \,\&\, \ldots \,\&\, \sigma(t_n) \overset{\sigma(t)}{\rightarrowtail} \tau(x_n), \\
 & & \textit{then } \tau \models_v^{\mathcal{M}} B
\end{array}$$

- A is *true at* w *in* $\mathcal{M}$, $\models_w^{\mathcal{M}} A$, iff for all w-assignment σ, $\sigma \models_w^{\mathcal{M}} A$.
- A is *true in* $\mathcal{M}$, $\models^{\mathcal{M}} A$, iff for all $w \in \mathcal{W}$, $\models_w^{\mathcal{M}} A$.
- A is *valid on* $\mathcal{F}$, $\mathcal{F} \models A$, iff for all models $\mathcal{M}$ based on $\mathcal{F}$, $\models^{\mathcal{M}} A$.

4.0.1. *Rigidity again*

An interpretation function I is said to satisfy the *rigidity condition* if

$$\text{if } d \in U_w \text{ and } d \prec v \text{ then} \qquad \begin{array}{l} I_w(c) \overset{d}{\rightarrowtail} I_v(c), \text{ and for all} \\ a_1, \ldots, a_n \in U_w \text{ and all } b_1, \ldots, b_n \in U_v, \\ \text{if } a_1 \overset{d}{\rightarrowtail} b_1 \text{ and } \ldots \text{and } a_n \overset{d}{\rightarrowtail} b_n, \text{ then} \\ (I_w(f^n))(a_1, ..., a_n) \overset{d}{\rightarrowtail} (I_v(f^n))(b_1, ..., b_n). \end{array}$$

- When the interpretation function satisfies the *rigidity condition* the following formula is valid for all terms $t, t_1 \ldots t_n$

$$\text{RG}_e \qquad |t : {}^{t_1}_{x_1} \ldots {}^{t_n}_{x_n}|A(x_1 \ldots x_n) \to |t : v_1 \ldots v_k|A[t_1/x_1 \ldots t_n/x_n]$$

where $v_1 \ldots v_k$ are the variables occurring in $t_1 \ldots t_n$

- When the rigidity condition is limited to some individual constant c and agent $a \in U_w$, then the following formula holds:

$$\text{RG}^{c,a}_e \qquad |t : x_1 \ldots x_n, {}^{c}_{x}|A(x_1 \ldots x_n, x) \to |t : x_1 \ldots x_n|A(x_1 \ldots x_n, c/x)$$

where $I_w(t) = a$.

Consider now the particular case of RG_e in which $t_1 \ldots t_n$ are variables and t is any term. Then we get

$$\text{RG}^v_e \qquad |t : {}^{y_1}_{x_1} \ldots {}^{y_n}_{x_n}|A(x_1 \ldots x_n) \to |t : v_1 \ldots v_k|A[y_1/x_1 \ldots y_n/x_n]$$

where $v_1 \ldots v_k$ are the variables $y_1 \ldots y_n$ without repetitions.

RG^v_e is valid, no matter if the interpretation function satisfies the rigidity condition or not. This is the reason why variables are said to be rigid designators. This terminology doesn't want to suggest any reference to the notion of rigidity in general.

- Let $a \in U_w$ and $a \prec v$. An individual constant c is said to be *a-stable* iff for all $e \in U_v$, $(I_w(c) \overset{a}{\rightarrowtail} e$ iff $e = I_v(c))$

When c is a-stable, the following formula holds:

$$\text{ST}^{c,a}_e \qquad |t : x_1 \ldots x_n, {}^{c}_{x}|A(x_1 \ldots x_n, x) \leftrightarrow |t : x_1 \ldots x_n|A(x_1 \ldots x_n, c/x)$$

where $I_w(t) = a$.

5. Correspondence

We said from the outset that we were looking for a general setting in which to envisage many different options by adding or removing some constraint. Here we show which principles correspond to which constraints on the structures, so according to the different applications one can choose the appropriate class of structures. First of all we consider constraints that capture some properties of knowledge that are relevant also at the propositional level. Then we consider constraints that capture properties that are specific of quantified epistemic logics.

Let an epistemic transition structure $\mathcal{F} = \langle \mathcal{W}, \mathcal{U}, \mathcal{D}, \prec, \rightarrowtail \rangle$ be given.

- $T_e : |x : \vec{y}|A \to A$
- $\mathcal{F}$ is reflexive iff (1) $\quad \forall a \in U_w(a \prec w)$ and (2) $\quad \forall a, d \in U_w(d \overset{a}{\rightarrowtail} d)$

Theorem 5.1. (Veridicity) *T_e is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is reflexive.*

Proof. $\Leftarrow$) By *reductio*. Let's assume there is a reflexive $\mathcal{F}$ s.t.:

(1) $\sigma \not\models_w |x : y|A \to A$
(2) $\sigma \models_w |x : y|A$
(3) $\sigma \not\models_w A$
(4) $\forall v \forall \tau(\sigma(x) \prec v \,\&\, \sigma(y) \overset{\sigma(x)}{\rightarrowtail} \tau(y) \Rightarrow \tau \models_v A)$

Given that $\mathcal{F}$ is reflexive, $\sigma(x) \prec w$ and $\sigma(y) \overset{\sigma(x)}{\rightarrowtail} \sigma(y)$, hence $\sigma \models_w A$, in contradiction with (3).

$\Rightarrow$) By contrapposition. Let $\mathcal{F}$ be non reflexive, so that (i) $\exists a \in U_w$ *not* $(a \prec w)$ or (ii) $\exists a, d \in U_w$ *not* $(d \overset{a}{\rightarrowtail} d)$. Let $\mathcal{I}$ be such that

$$\text{for all } v \in \mathcal{W}, I_v(P) = \{b \in U_v : b \neq d\}$$

If *not* $(a \prec w)$, we have that $\sigma^{x \triangleright a} \models_w |x : \star|\forall y P(y)$, and $\sigma^{x \triangleright a} \not\models_w \forall y P(y)$. If *not* $(d \overset{a}{\rightarrowtail} d)$, then $\sigma^{x \triangleright a,\, y \triangleright d} \models_w |x : y|P(y)$, and $\sigma^{x \triangleright a,\, y \triangleright d} \not\models_w P(y)$. Either way there is a counterexample to T_e.[22]

■

- $D_e : |x : \vec{y}|A \to < x : \vec{y} > A$
- $\mathcal{F}$ is serial iff (1) $\quad \forall a \in U_w \exists v(a \prec v)$ and (2)$\forall a, b \in U_w \exists c \in U_v(b \overset{a}{\rightarrowtail} c)$

Theorem 5.2. (Consistency) *D_e is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is serial.*

- $4_e : |x : \vec{y}|A \to |x : x\,\vec{y}|\,|x : \vec{y}|A$

[22] Note that if we impose on $\mathcal{F}$ only the condition (1), but not (2), we have a structure where are valid only the instances of T_e where A is a closed formula, thus validating T_e for *de dicto* knowledge, but not for *de re* knowledge. The same point holds for the other propositional properties of knowledge.

- $\mathcal{F}$ is transitive iff
  (1) $\forall a \in U_w, d \in U_v (a \overset{a}{\rightarrowtail} d \,\&\, d \prec z \Rightarrow a \prec z)$
  and
  (2) $\forall a, b \in U_w,\ c, d \in U_v,\ e \in U_z (a \overset{a}{\rightarrowtail} d \,\&\, b \overset{a}{\rightarrowtail} c \,\&\, c \overset{d}{\rightarrowtail} e \Rightarrow a \overset{a}{\rightarrowtail} e)$

Theorem 5.3. (Positive introspection) 4_e *is valid on* $\mathcal{F}$ *if and only if* $\mathcal{F}$ *is transitive.*

- $5_e : \neg|x : \vec{y}|A \to |x : x\vec{y}|\,\neg|x : \vec{y}|A$

- $\mathcal{F}$ is Euclidean iff
  (1) $\forall a \in U_w,\ b \in U_v (a \prec v \,\&\, a \prec z \,\&\, a \overset{a}{\rightarrowtail} b \Rightarrow b \prec z)$
  and
  (2) $\forall a, d \in U_w,\ \ c, b \in U_v,\ e \in U_z\ (d \overset{a}{\rightarrowtail} c \,\&\, d \overset{a}{\rightarrowtail} e \,\&\, a \overset{a}{\rightarrowtail} b \Rightarrow c \overset{b}{\rightarrowtail} e)$

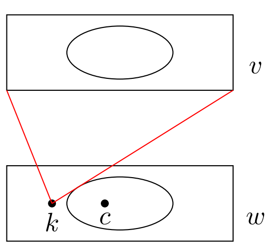
Theorem 5.4. (Negative introspection) 5_e *is valid on* $\mathcal{F}$ *if and only if* $\mathcal{F}$ *is Euclidean.*

- $GF : \exists y|t : x_1 ... x_n\, y|A \to |t : x_1 ... x_n|\exists y A$

- $\mathcal{F}$ is $\mathcal{D}$-totally defined iff $\forall k \in U_w (k \prec v \Rightarrow \forall a \in D_w \exists b \in D_v (a \overset{k}{\rightarrowtail} b))$

Theorem 5.5. GF *is valid on* $\mathcal{F}$ *if and only if* $\mathcal{F}$ *is* $\mathcal{D}$*-totally defined.*

Proof. $\mathcal{F}$ not $\mathcal{D}$-totally defined implies $\mathcal{F} \not\models GF$



$$I_v(P) \cap D_v = \emptyset$$

$$\sigma^{x \rhd c} \models_w |k : x|P(x)$$

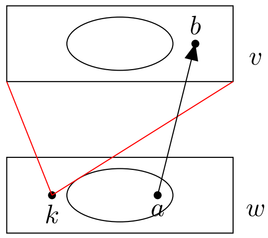
$$\sigma^{x \rhd c} \not\models_w |k : \star|\exists x P(x)$$

∎

- $CBF : |t : x_1 ... x_n|\forall y A \to \forall y|t : x_1 ... x_n\, y|A$

- $\mathcal{F}$ is $\mathcal{D}$-preservative iff $\forall k \in U_w (k \prec v \Rightarrow \forall a \in D_w \forall b \in U_v (a \overset{k}{\rightarrowtail} b \Rightarrow b \in D_v))$

Theorem 5.6. CBF *is valid on* $\mathcal{F}$ *if and only if* $\mathcal{F}$ *is* $\mathcal{D}$*-preservative.*



$$I_v(P) = D_v$$

$$\sigma^{x \triangleright a} \models_w |k : \star|\forall x P(x)$$

$$\sigma^{x \triangleright a} \not\models_w |k : x|P(x)$$

$$\sigma^{x \triangleright a} \not\models_w |k : \star|\forall x P(x) \to \forall x|k : x|P(x)$$

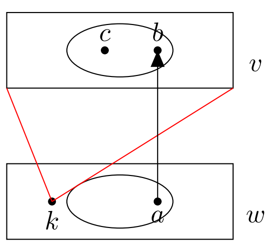
PROOF. $\mathcal{F}$ not $\mathcal{D}$-preservative implies $\mathcal{F} \not\models CBF$.

■

- $BF : \forall y|t : x_1...x_n\, y|A \to |t : x_1...x_n|\forall y A$
- $\mathcal{F}$ is $\mathcal{D}$-surjective iff $\forall k \in U_w(k \prec v \Rightarrow \forall b \in D_v \exists a \in D_w(a \overset{k}{\rightarrowtail} b))$

THEOREM 5.7. *BF is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is $\mathcal{D}$-surjective.*

PROOF. $\mathcal{F}$ not $\mathcal{D}$-surjective implies $\mathcal{F} \not\models BF$.



$$D_w = \{a\} \qquad b \in I_v(P) \qquad c \notin I_v(P)$$

$$\sigma \models_w \forall x|k : x|P(x)$$

$$\tau \not\models_v \forall x P(x)$$

$$\sigma \not\models_w \forall x|k : x|P(x) \to |k : \star|\forall x P(x)$$

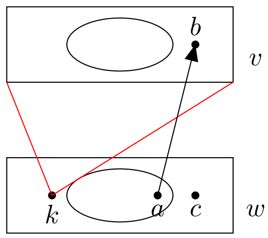
■

- $SHRT : |t : x_1...x_n\, y|A \to |t : x_1...x_n|A$
- $\mathcal{F}$ is $\mathcal{U}$-totally defined iff $\forall k \in U_w(k \prec v \Rightarrow \forall a \in U_w \exists b \in U_v(a \overset{k}{\rightarrowtail} b))$

THEOREM 5.8. *$SHRT$ is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is $\mathcal{U}$-totally defined.*

PROOF. $\mathcal{F}$ not $\mathcal{U}$-totally defined implies $\mathcal{F} \not\models SHRT$.

■



$$b \notin I_v(P)$$

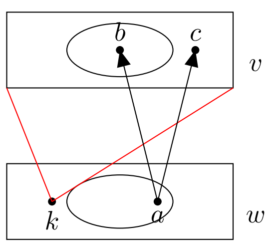
$$\sigma^{x \triangleright c, y \triangleright a} \models_w |k : x\, y| P(y)$$

$$\sigma^{x \triangleright c, y \triangleright a} \not\models_v |k : y| P(y)$$

- $NI :\ x = y \to |t : x\, y| x = y$
- $\mathcal{F}$ is $\mathcal{U}$-functional iff $\forall k \in U_w (k \prec v \Rightarrow \forall a \in U_w \forall b, c \in U_v (a \overset{k}{\rightarrowtail} b \wedge a \overset{k}{\rightarrowtail} c \Rightarrow b = c))$

THEOREM 5.9. *NI is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is $\mathcal{U}$-functional.*

PROOF. $\mathcal{F}$ not $\mathcal{U}$-functional implies $\mathcal{F} \not\models NI$.



$$c \neq b$$

$$\sigma^{x \triangleright a, y \triangleright a} \models_w x = y$$

$$\tau^{x \triangleright c, y \triangleright b} \not\models_w x = y$$

$$\sigma^{x \triangleright a, y \triangleright a} \not\models_w x = y \to |k : x\, y| x = y$$

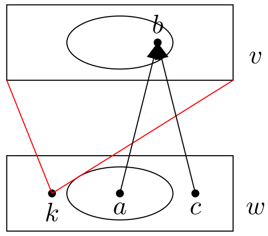
■

- $ND :\ x \neq y \to |t : x\, y| x \neq y$
- $\mathcal{F}$ is $\mathcal{U}$-not-convergent iff $\forall k \in U_w (k \prec v \Rightarrow \forall a, b \in U_w \forall c \in U_v (a \overset{k}{\rightarrowtail} c \wedge b \overset{k}{\rightarrowtail} c \Rightarrow a = b))$

THEOREM 5.10. *ND is valid on $\mathcal{F}$ if and only if $\mathcal{F}$ is $\mathcal{U}$-not-convergent.*

PROOF. $\mathcal{F}$ not $\mathcal{U}$-not-convergent implies $\mathcal{F} \not\models ND$.

■



$$b = b$$

$$\sigma^{x \rhd a, y \rhd c} \models_w x \neq y$$

$$\tau^{x \rhd b, y \rhd b} \not\models_w x \neq y$$

$$\sigma^{x \rhd a, y \rhd c} \not\models_w x \neq y \to |k : x\, y|x \neq y$$

6. Axiomatic system

In [5] is given a sound and complete axiomatization for the logic $Q.K_e$ of the class of all structures such that, for all $w \in \mathcal{W}$, $U_w = D_w$.

In this section we present a sound and complete axiomatization for the logic $Q^0.K_e$ of the class of all transition frames. For the sake of simplicity, we introduce the existence predicate $\mathcal{E}(x)$ with the usual definition $\mathcal{E}(x) =_{df} \exists y(x = y)$. The logic $Q^0.K_e$ is so axiomatized:

- **Axioms**

PC A basis for propositional tautologies.

PRM_e $|x : x_1 \ldots x_n|A \leftrightarrow |x : x_{i_1} \ldots x_{i_n}|A$
for any permutation $x_{i_1} \ldots x_{i_n}$ of $x_1 \ldots x_n$

K_e $|x : \vec{y}|(A \to B) \to (|x : \vec{y}|A \to |x : \vec{y}|B)$

LNGT_e $|x : x_1 \ldots x_n|A \to |x : x_1 \ldots x_n, x_{n+1}|A$

RG_e^v $|x : {}^{\vec{y}}_{\vec{z}}|A \to |x : y_1 \ldots y_k|(A[\vec{y}/\vec{z}])$
where $y_1 \ldots y_k$ are all the different variables in $\vec{y}$

ID $x = x$

LBZ $s = t \to (A[s/x] \to A[t/x])$

$\mathcal{E}$I $\forall x A \to (\mathcal{E}(t) \to A[t/x])$

- **Rules**

MP $$\frac{A \quad A \to B}{B}$$

N $$\frac{A}{|x : x_1 \ldots x_n|A}$$ provided $fv(A) \subseteq \{x_1, \ldots, x_n\}$

$$\text{SFV} \qquad \frac{A}{A[t/x]} \qquad \text{provided } t \text{ is free for } x \text{ in } A$$

$$\mathcal{E}G \qquad \frac{A \to (\mathcal{E}(t) \to B[t/x])}{A \to \forall x B} \quad x \text{ not free in } A$$

Basically the axiomatization of $Q^0.K_e$ is the axiomatization for $Q.K_e$ presented in [5] where UI and UG are substituted by their restricted version $\mathcal{E}I$ and $\mathcal{E}G$. Soundness and completeness can be proved along the lines of the proof in [5].

7. Why index-modalities are better than λ-operator

One may wonder whether the fine-grained analysis of the interplay between modalities and quantifiers due, in our opinion, to the language with indexed modalities and shown in the correspondence theorems, could be achieved by adding the λ-operator to a term-modal language. The answer is in the negative, since, by so doing, there are formulas that become provable whereas they should not; the presence of the λ-operator doesn't make any difference. As a paradigmatic case consider the formula $\exists x|t|Px \to |t|\exists xPx$ and its proof in the sequent calculus for K as presented in [9]: 144.

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{}{P(p), \neg P(p), \forall x \neg P(x)}\ \text{ax}}{P(p), \forall x \neg P(x)}\ \forall}{|t|P(p), \langle t\rangle \forall x \neg P(x)}\ \langle t\rangle}{\exists x|t|P(x), \langle t\rangle \forall x \neg P(x)}\ \exists}{\exists x|t|P(x) \wedge \langle t\rangle \forall x \neg P(x)}\ \wedge$$

The validity of $\exists x|t|Px \to |t|\exists xPx$ is rather questionable and yet its proof seems to use unquestionable inference rules. Not only that, but that proof would remain untouched by the presence in the language of the λ-operator. Instead, if we rephrase the formula $\exists x|t|Px \to |t|\exists xPx$ in our language with indexed modalities we get $\exists x|t : x|Px \to |t : \star|\exists xPx$, *i.e.* an instance of the schema GF. As shown in theorem 5.5, GF corresponds to a specific condition on the transition relations and it should not be provable in the minimal quantified extension of K.

Let us redo the above proof in our language, starting from below:

$$
\dfrac{\dfrac{\dfrac{????}{|t : \genfrac{}{}{0pt}{}{p}{x}|P(x), \langle t : \star\rangle \forall x \neg P(x)}\ ???}{\exists x |t : x| P(x), \langle t : \star\rangle \forall x \neg P(x)}\ \exists}{\exists x |t : x| P(x) \wedge \langle t : \star\rangle \forall x \neg P(x)}\ \wedge
$$

The modal operators of the formulas of the top sequent differ as far as the indices are concerned and so a rule analogous to the rule $\langle t\rangle$ cannot be applied and the sequent $P(x), \forall x \neg P(x)$ cannot be obtained. What is needed is a new rule, R_{SHRT}, corresponding to the principle of shortening, SHRT, according to which indices not occurring in the formula that follows the box-operator can be deleted:

SHRT: $\qquad |t : x_1 \ldots x_n, z|A \rightarrow |t : x_1 \ldots x_n|\, A$

R_{SHRT} can be formulated as follows in analogy with the rules of [9]: 144.

$$
\dfrac{S, \langle t : \genfrac{}{}{0pt}{}{\vec{s}\ p}{\vec{x}\ y}\rangle A}{S, \langle t : \genfrac{}{}{0pt}{}{\vec{s}}{\vec{x}}\rangle A}\ R_{SHRT}
$$

Now the proof can go through:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P(x), \neg P(x), \forall x \neg P(x)}\ \text{ax}}{P(x), \forall x \neg P(x)}\ \forall}{|t : \genfrac{}{}{0pt}{}{p}{x}|P(x), \langle t : \genfrac{}{}{0pt}{}{p}{x}\rangle \forall x \neg P(x)}\ \langle t : \genfrac{}{}{0pt}{}{p}{x}\rangle}{|t : \genfrac{}{}{0pt}{}{p}{x}|P(x), \langle t : \star\rangle \forall x \neg P(x)}\ R_{SHRT}}{\exists x |t : x| P(x), \langle t : \star\rangle \forall x \neg P(x)}\ \exists}{\exists x |t : x| P(x) \wedge \langle t : \star\rangle \forall x \neg P(x)}\ \wedge
$$

Where the problem lies is now clear: the proof of $\exists x|t : x|Px \rightarrow |t : \star|\exists x Px$ involves one step (R_{SHRT}) that goes unnoticed in the proof of [9] just because of the language used. This step is valid only by assuming that every object can be "retraced" in any accessible world; in our terminology, R_{SHRT} is valid if the transition relation is everywhere defined on the outer domains.

We conclude by noticing that GF and $SHRT$ are provably equivalent in a quantified epistemic logic based on classical logic, see [4] and [5]. Here is a proof of an instance of SHRT, $|t : x, y|P(x) \rightarrow |t : x|P(x)$, by the help of R_{GF} and the cut rule.

$$
\dfrac{
\dfrac{\dfrac{}{|t : x\,y|P(x),\, \neg|t : x\,y|P(x)}\,\text{ax}}{|t : x\,y|P(x),\, \forall y \neg|t : x\,y|P(x)}\,\forall
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{P(x),\, \neg P(x)}\,\text{ax}}{\exists y P(x),\, \neg P(x)}\,\text{vacuous quant.}}{|t : x|\exists y P(x),\, \langle t : x\rangle \neg P(x)}\,\langle t : x\rangle}{\exists y|t : x\,y|P(x),\, \langle t : x\rangle \neg P(x)}\,R_{GF}}{}
}{
\dfrac{|t : x\,y|P(x),\, \langle t : x\rangle \neg P(x)}{|t : x\,y|P(x) \wedge \langle t : x\rangle \neg P(x)}\,\wedge
}\,\text{cut rule}
$$

A proof of $|t : x, y|P(x) \to |t : x|P(x)$ enjoying the subformula property seems hard to obtain, and consequently a cut free sequent calculus of the classically quantified epistemic logic $Q_e.K + GF$, see [5]. A further motivation in favour of a free quantified epistemic logic can be retraced in the fact that the schemata GF and $SHRT$ are not provably equivalent.

References

[1] Burge, T., 'Belief *De Re*', *The Journal of Philosophy* 74: 338–362, 1977.

[2] Carlson, L., 'Quantified Hintikka-Style Epistemic Logic', *Synthese* 74: 223–262, 1998.

[3] Colombetti, M. and A. Lomuscio, 'QLB: A Quantified Logic of Belief', in Müller, Woolridge, and Jennings (eds.), *Intelligent Agents III*, Springer-Verlag, 1997.

[4] Corsi, G., 'Necessary for', in Glymour, Wei, and Westerstahl (eds.), *Logic, Methodology and Philosophy of Science, Proceedings of the Thirtheen International Congress*, King's College Publications, 2009.

[5] Corsi, G., 'A New Approach to Epistemic Logic', *forthcoming.*

[6] Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi, *Reasoning about Knowledge*, MIT Press, 1995.

[7] Fitting, M. C., 'Modal Logic Should Say More than it Does', in Lassez, and Plokin (eds.), *Computational Logic*, MIT Press, 1991.

[8] Fitting, M. C., and R. L. Mendelsohn, *First-Order Modal logic*, Kluwer,1998.

[9] Fitting, M. C, L. Thalmann, and A. Voronkov, 'Term-Modal Logic', *Studia Logica*, 69: 133–169, 2001.

[10] Gochet, P., and P. Gribomont, 'Epistemic Logic', in Gabbay, and Wood (eds.), *The Handbook of History and Philosophy of Logic, vol 6*, Elsevier, 2006.

[11] GROVE, A. J.,'Naming and Identity in Epistemic Logic part II: a First-Order Logic for Naming', *Artificial Intelligence* 74: 31–350, 1995.

[12] HEIM, I.,'Presupposition Projection and the Semantics of Attitudes Verbs', *Journal of Semantics* 9: 183–221, 1992.

[13] HENDRICKS, V. F., K. F. JORGENSEN, and S. A. PEDERSEN, 'Introduction', in Hendricks, Jorgensen, and Pedersen (eds.), *Knowledge Contributors*, Kluwer, 2003.

[14] HINTIKKA, J., *Knowledge and Belief*. Cornell University Press, 1962.

[15] HINTIKKA, J., 'A Second Generation Epistemic Logic and Its General Significance', in Hendricks, Jorgensen, and Pedersen (eds.), *Knowledge Contributors*, Kluwer, 2003.

[16] KAPLAN, D., 'Quantifying in', *Synthese* 19: 178–214, 1968.

[17] LEWIS, D., 'Attitudes *De Dicto* and *De Se*', *The Philosophical Review* 88: 513–543, 1979.

[18] MEYER, J-J. Ch., and W. VAN DER HOEK, *Epistemic logic for AI and Computer Science*, Cambridge University Press, 1995.

[19] QUINE, W. V. O., 'Quantifiers and Propositional Attitudes', *The Journal of Philosophy* 53: 177–187, 1956.

[20] VON WRIGHT, G. H., *An Essay in Modal Logic*, North-Holland, 1951.

GIOVANNA CORSI
Dipartimento di Filosofia
Università di Bologna
via Zamboni 38 I-40126 Bologna
giovanna.corsi@unibo.it

EUGENIO ORLANDELLI
Dipartimento di Filosofia
Università di Bologna
via Zamboni 38 I-40126 Bologna
eugenio.orlandelli@studio.unibo.it