

# STEPS TOWARDS A FIRST-ORDER LOGIC OF EXPLICIT AND IMPLICIT BELIEF

Gerhard Lakemeyer  
Department of Computer Science  
University of Toronto  
Toronto, Ontario  
Canada, M5S 1A4

## Abstract

Modelling the beliefs of an agent who lacks logical omniscience has been a major concern recently. While most of the work has concentrated on propositional logics of belief, this paper primarily addresses issues raised by adding quantifiers to such logics. In particular, we are focusing on *quantifying in* and the distinction between “knowing what” and “knowing that”. After arguing why a model of limited reasoning should preserve this distinction, we show how this can be accomplished by a semantics based on a restricted form of tautological entailment.

## Introduction

Ever since *possible-worlds semantics* was first proposed for models of knowledge and belief, [Hint62]<sup>1</sup> it has been argued that this semantics is not a realistic model for reasoning, mainly because it requires the reasoner to be *logically omniscient*, that is, beliefs are closed under logical implication. Clearly, the assumption that, among other things, an agent knows *all* valid sentences is quite unnatural and it also leads to serious computational problems. We adopt the view of [Leve84b] and [FaHa85] that possible-worlds semantics is good for expressing what is *implicit* in an agent's knowledge, but more restricted models should be used to characterize what is believed *explicitly*. As a formal tool in the study of belief, these authors use modal logics, an approach followed also in this work.

An interesting application of belief logics is the study of knowledge bases (KBs) that are part of larger systems requiring quick responses from their KBs, as in robots. If the KB is characterized as a finite set of sentences in a language like first-order logic, fast limited reasoning may be favored over full theorem proving. The efficiency versus completeness trade-offs are only some of many issues which can be studied very elegantly using belief logics, as demonstrated in [Leve84b] and [FaHa85].

The literature on limited reasoning reveals that there are at least two major approaches to avoiding logical omniscience. One is syntactic in nature; the other tries to give a semantic account. An example of the former is to limit the set of inferences by describing models of belief as explicit sets of sentences, as is done in [Eber74] and [MoHe79]. Alternatively, [Kono84] models beliefs using a base set of sentences together with a possibly incomplete set of deduction rules. One fundamental problem with this syntactic approach is that the answer to the question of which sentences should be believed can be quite arbitrary and lacks an intuitive semantic account.

Levesque [Leve84b] is among the first who attempt to address exactly this issue and give a plausible semantics for explicit beliefs. He restricts himself to a propositional logic of explicit and implicit belief. He uses two modal operators  $B$  and  $L$  to capture this distinction, a convention we will also follow in this paper. A sentence  $B\alpha$  is then said to be true if  $\alpha$  is explicitly believed by the agent. Similarly,  $L\alpha$  expresses that  $\alpha$  is believed implicitly. The key idea for his semantics of explicit belief is that worlds (or *situations*, as he calls them) may be incompletely specified or even contradictory. This is realized in the following way: for any situation  $s$  and propositional letter  $p$ ,  $s$  will support the truth of  $p$ , its falsity, both or none, which amounts to assigning  $p$  one of four truth values  $\{\}, \{t\}, \{f\}, \{t, f\}$  instead of the usual two. A sentence  $\alpha$  is then said to be explicitly believed if the truth of  $\alpha$  is supported in all situations the agent thinks possible (the *belief set*). This avoids logical omniscience in that the truth of a sentence no longer bears any relationship to its negation, which is crucial for *modus ponens*. At the same time, contradictory beliefs like  $B(p \wedge \neg p)$  are possible. Levesque shows that his semantics for explicit beliefs coincides with that of tautological entailment in relevance logic (as in [AnBe75] and [Beln77]).

---

<sup>1</sup>Originally, Hintikka developed his semantics using so-called *model sets*, but he later [Hint71] recast the theory in terms of the now standard notions of possible-worlds semantics.

In the same spirit, Fagin and Halpern [FaHa85] propose several propositional logics for belief that try to give plausible accounts of why people are not logically omniscient. Examples are lack of awareness, resource-boundedness, not knowing the relevant rules, and focus of attention. In contrast to Levesque, the authors stick to two-valued logics and allow for multiple agents. Being two-valued means that contradictory beliefs of the form  $B(p \wedge \neg p)$  can no longer be satisfied. However, their “society of minds” logic allows sentences of the form  $Bp \wedge B\neg p$  to be satisfiable (see also [Leve] for a similar proposal). Finally, Fagin and Halpern suggest treatments for nested beliefs, which Levesque did not consider in his original paper.

Given that all of the above was based on propositional logic, the obvious question arises as to whether we can find similar quantified versions. That is a non-trivial matter. A major stumbling block is that first-order logic by itself is undecidable. Patel-Schneider [Pate85] proves that an existing model of limited reasoning, namely the first-order version of tautological entailment, is also undecidable. By modifying this model, however, he arrives at a decidable version. (See also [Fris85], who uses quite different methods with surprisingly similar results). While this seems to be a good starting point for investigating models of explicit belief — and in fact, that is exactly what we are going to propose — by itself it is not entirely satisfactory. Since Patel-Schneider’s language is that of classical first-order logic, it lacks the primitives to talk about beliefs (such as the  $B$  or  $L$  operator). Thus it cannot address important issues like nested beliefs or quantifying in.

In this paper we will concentrate on the problem of *quantifying in* with respect to a first-order logic with modal operators  $B$  and  $L$  assuming a single agent. We start with a brief discussion of the distinctions between “knowing that” and “knowing what”. We give reasons why we think those distinctions should not be given up in the context of explicit belief. After introducing Patel-Schneider’s model of limited reasoning, we show that its obvious extension to a belief logic has some rather undesirable properties. We then develop a model that solves these problems and maintains the key features of quantifying in.

## “Knowing What” and “Knowing That”

Since Frege, the phenomenon that Leibniz’ principle of substituting equals for equals seems to break down in propositional attitudes like *knowing* has been of considerable concern among philosophers (for example, [Freg92] and [Lins71]). One manifestation of the problem is the distinction between “knowing what” and “knowing that”.<sup>2</sup> For example, I may know that the current President of the US lives in the White House without knowing who he (or she!) is.

Belief logics that want to be able to distinguish between the two cases usually allow two ways of denoting individuals. One uses *standard names* or *rigid designators*, where there can be no doubt about the identity of the referred object; e.g., “7” always denotes one particular number. The other uses *non-rigid designators*, where the individual referred to may vary from interpretation to interpretation like the term “President of the US”.

---

<sup>2</sup>The main distinction made in logic between knowledge and belief is that knowledge must be true in the actual world, while beliefs may not. Although we favor the latter, this distinction is of no consequence for this work and we sometimes even use “know” instead of “believe”.

These two classes of names plus the distinction between quantifiers outside and inside the scope of belief operators make it possible to capture some of the intuition behind “knowing what” and “knowing that”. In logics of implicit beliefs that are based on classical modal logics like S5 or weak S5 (as in [Moor80] or [Leve82]), the following sentences come out valid ( $\models$ ) or not valid ( $\not\models$ ) in a possible worlds semantics with a fixed universe of discourse over all worlds: Let  $P$  be a unary predicate,  $x$  a variable,  $a$  a non-rigid designator, and  $n$  a standard name.

- 1)  $\models \exists xLPx \supset L\exists xPx$
- 2)  $\not\models L\exists xPx \supset \exists xLPx$
- 3)  $\models LPa \supset L\exists xPx$
- 4)  $\not\models LPa \supset \exists xLPx$
- 5)  $\models LPn \supset \exists xLPx$
- 6)  $\models L\forall xPx \supset \forall xLPx$
- 7)  $\models \forall xLPx \supset L\forall xPx$

As an example, if we read  $L$  as “is known” and  $P$  as “is a teacher”, the first two statements reflect the distinction between “knowing what” and “knowing that” as follows: The first says that whenever some particular individual (identified by his standard name) is known to be a teacher, then it is known that there exists a teacher. The second, on the other hand, states that knowing that there is a teacher does not imply that there is a particular individual who is known to be a teacher.

Assuming that the above seven statements are reasonable choices for implicit beliefs, the question arises whether a model of explicit belief can or should have a different view. To us, tampering with 1 or 3 does not make much sense. Falsifying 1, for example, amounts to allowing for an agent who believes  $P$  for a known individual, but fails to acknowledge the existence of one!

The most important point we want to make, however, is that 2 and 4 should not come out valid under explicit belief. Doing so would essentially collapse the distinction between “knowing that” and “knowing what”, which has a grave consequence for the expressiveness of the representation language. There would be no way of telling a knowledge base that somebody is a  $P$  without telling it who it is! At best, incomplete knowledge of this kind could be approximated by finite disjunctions, but there would be no existential quantification.

Finally, let us take a look at the last sentence,  $\forall xLPx \supset L\forall xPx$ . Essentially, it expresses that the rule of *universal generalization* applies to implicit belief. We feel that this may not necessarily be the case in explicit beliefs.<sup>3</sup> Since in propositional models of explicit belief *modus ponens* had to be given up as a valid inference rule within beliefs, it should not be too surprising if universal generalization fails when adding quantifiers.

<sup>3</sup>Even in possible worlds semantics this sentence, the controversial *Barcan Formula*, can fail, if worlds are allowed to have differing universes of discourse [HuCr68].

The main goal of the rest of the paper is to develop a suitable semantics for explicit belief. It will have the properties 1-6 concerning quantifying in, but not 7 (universal generalization fails). Implicit beliefs, which will be modelled in the standard possible worlds fashion, will of course exhibit all seven properties.

## The Language $\mathcal{L}$

The language we use is basically a straightforward extension of Levesque's propositional logic of explicit and implicit beliefs [Leve84b] to the first-order case. The only non-standard feature is the explicit representation of standard names, which are called *parameters* in the language.

$\mathcal{L}$  consists of countably infinite sets of *variables*  $V$ , *parameters*  $N$ , and *function and predicate symbols*. Function symbols with no arguments are also called *constants*, but note that they are non-rigid designators and not to be confused with parameters, the rigid designators of our language. A *term* is either a variable, a parameter, or a function symbol whose arguments are themselves terms. A *closed term* is a term not containing any variables.  $\mathcal{T}$  denotes the set of all closed terms. An *atomic formula* (or simply *atom*) is a predicate symbol with terms as arguments. Given the atomic formulas, the logical connectives  $\neg$  and  $\vee$ , the existential quantifier  $\exists$ , and modal operator symbols  $B$  and  $L$ , we can generate all well-formed formulas of  $\mathcal{L}$  using the standard formation rules.<sup>4</sup> Given the usual meaning of *bound* and *free variables*, we say  $\alpha^{x/t}$  is that formula obtained from  $\alpha$ , where all occurrences of the free variable  $x$  are replaced by the term  $t$ . The same convention is used if  $\alpha$  is a term instead of a formula. Finally, a formula with no free variables is called a *sentence*.

As noted at the beginning of this section, parameters serve as standard names or rigid designators in our language. We refer to them as numerals  $1, 2, 3, \dots$ . They are needed, as pointed out in the previous section, so that we can distinguish "knowing what" from "knowing that". In order to give them their intended meaning, we will simply use the parameters themselves as the fixed universe of discourse in all models. For a more thorough discussion of why this trick is sound, and for more reasons why the use of standard names comes in handy in modal belief logics, see [Leve82].

Since we are not concerned with nested beliefs in this discussion, we restrict  $\mathcal{L}$  to that set of formulas that have no nested occurrences of  $B$  or  $L$ . Furthermore, we write  $\mathcal{L}_1$  for the subset of  $\mathcal{L}$  with *no* occurrences of  $B$  or  $L$ .

## t-Entailment

The notion of explicit belief that we will put forward in this paper has a strong resemblance to Patel-Schneider's notion of *t-entailment*, a variant of tautological entailment in first-order relevance logic. The semantics for this decidable form of entailment uses world descriptions that are straightforward generalizations of situations introduced in [Leve84b]: for any predicate symbol  $P$  with parameters  $(\vec{n})$  as arguments, a situation  $s$  supports either the truth of  $P\vec{n}$ ,

<sup>4</sup>Other connectives like  $\wedge$ ,  $\supset$ , and the universal quantifier  $\forall$  are defined in the usual way from  $\neg$ ,  $\vee$ , and  $\exists$ .

its falsity, both, or none. In addition, every closed term is mapped into a standard name (its meaning with respect to  $s$ ) using a *coreference relation* as in [Leve84a]. A coreference relation  $\equiv$  can be characterized as follows:

- a)  $\equiv$  is an equivalence relation.
- b) No two standard names corefer.
- c) Every closed term has a coreferring standard name.
- d) If  $t_1$  and  $t_2$  corefer, so do  $t^{x/t_1}$  and  $t^{x/t_2}$ , where  $t$  contains one free variable  $x$ .

It is easy to see that every closed term has a unique coreferring standard name. A situation  $s$  is a triple

$$s = \langle T^s, F^s, \equiv_s \rangle .$$

$T^s$  and  $F^s$  map every  $k$ -ary predicate symbol into a  $k$ -place relation over  $N$ . Truth- and false-support of atomic formulas containing parameters corresponds to membership in either  $T^s$  or  $F^s$ .  $\equiv_s$  denotes the coreference relation of  $s$ .

We call the set of all situations  $S$ . We also use the following shorthand: given a situation  $s$ , a  $k$ -ary predicate symbol  $P$  and a vector of closed terms  $\vec{t} = t_1, \dots, t_k$ , we say  $(\vec{t}) \in T^s(P)$  if  $(n_1, \dots, n_k) \in T^s(P)$ , where  $n_i \equiv_s t_i$  (similarly for  $(\vec{t}) \in F^s(P)$ ).

Finally, in order to interpret formulas with variables, we introduce *variable maps*  $\nu$  from  $V$  into  $N$ .  $\mathcal{V}$  is the set of all such variable maps. We notate  $\nu_n^x$  as that variable map identical to  $\nu$  except that  $x$  gets mapped into  $n$ . For convenience we take the liberty of applying variable maps also to terms and sequences of terms with the intent that all occurrences of variables in a term become replaced by the standard name called for by the variable map (e.g.,  $\nu(f(x), g(y, x)) = f(3), g(6, 3)$  if  $\nu(x) = 3$  and  $\nu(y) = 6$ ).

We are now just about ready to present Patel-Schneider's semantics. A pivotal role in attaining a decidable form of entailment is played by his very restricted interpretation of existential quantification. Formulas are interpreted with respect to a *set* of situations that share a common coreference relation (called a *compatible* set of situations). Such a set  $S$  supports the truth of a formula  $\exists x \alpha$  only if there is a standard name  $n$  so that  $\alpha^{x/n}$  has true support in *every* element of  $S$ . We will refer to this way of interpreting  $\exists$  as *global existential quantification*. The semantics from [Pate85], with adaptations to our notation,<sup>5</sup> is: let  $S$  be a compatible set of situations,  $\nu$  a variable map. The language is  $\mathcal{L}_1$ , since this semantics does not account for belief operators.

$$1. S, \nu \models_T P\vec{t} \iff \text{for all } s \in S \ \nu(\vec{t}) \in T^s(P)$$

$$S, \nu \models_F P\vec{t} \iff \text{for all } s \in S \ \nu(\vec{t}) \in F^s(P)$$

$$2. S, \nu \models_T \neg \alpha \iff S, \nu \models_F \alpha$$

<sup>5</sup>Actually, it is a little more than just a notational change, since the language Patel-Schneider uses has no parameters. Adding them, however, is of no consequence for his results.

$$S, \nu \models_F \neg \alpha \iff S, \nu \models_T \alpha$$

$$3. S, \nu \models_T \alpha \vee \beta \iff \exists S_1 S_2 S = S_1 \cup S_2 \text{ s.t. } S_1, \nu \models_T \alpha \text{ and } S_2, \nu \models_T \beta$$

$$S, \nu \models_F \alpha \vee \beta \iff S, \nu \models_F \alpha \text{ and } S, \nu \models_F \beta$$

$$4. S, \nu \models_T \exists x \alpha \iff \text{for some } n \in N \ S, \nu_n^x \models_T \alpha$$

$$S, \nu \models_F \exists x \alpha \iff \text{for all } n \in N \ S, \nu_n^x \models_F \alpha$$

t-entailment ( $\longrightarrow_t$ ) is then defined as:  $\alpha \longrightarrow_t \beta$  if and only if for all compatible sets of situations  $S$  and variable maps  $\nu$ , if  $S, \nu \models_T \alpha$  then  $S, \nu \models_T \beta$ .

At first glance it seems that it is not  $\exists$  but rather  $\forall$  that gets a special treatment in this semantics. On closer inspection, however, it becomes clear that it is the special way of looking at existential quantification that causes the somewhat odd interpretation of  $\forall$ . Let's look at the example  $S, \nu \models_T (\exists x Px \vee \exists y Qy)$ . According to the semantics, we can partition  $S$  into  $S_1$  and  $S_2$  such that all of  $S_1$  supports  $Pn$  and all of  $S_2$  supports  $Qm$  for some fixed  $n$  and  $m$ . But that is just another way of saying: for some fixed parameters  $n$  and  $m$ ,  $Pn \vee Qm$  is supported in all situations of  $S$ , where  $\forall$  is now interpreted in the usual way.

One important property of t-entailment that seems to be crucial for it being decidable is the fact that  $(Pa \vee Pb) \not\longrightarrow_t \exists x Px$ . This is caused exactly by the special treatment of  $\exists$  we have just discussed. We feel that this is not an unreasonable limitation, since it reflects a certain constructive or intuitionistic mode of reasoning. Applying this feature to a logic of explicit and implicit belief would mean: while the existential is clearly *implicit* in the disjunction, we may not want to require that any agent who believes the disjunction explicitly automatically be able to reason by cases and come up with the existential.

## From t-entailment to Explicit Belief

Levesque based his semantics for a propositional belief logic on world descriptions that are situations, belief sets that are sets of situations, and tautological entailment. An analogous and probably the simplest extension to Patel-Schneider's semantics would have us view world descriptions as compatible sets of situations and a belief set as a set of compatible sets of situations. With this picture in mind, we get the following definition for explicit belief:

Let  $\Gamma$  be a set of compatible sets of situations, and let  $S$  and  $\nu$  be as before.

$$5. S, \nu \models_T B\alpha \iff \forall S' \in \Gamma \ S', \nu \models_T \alpha$$

$$S, \nu \models_F B\alpha \iff S, \nu \not\models_T B\alpha$$

Very soon one discovers that this approach has some rather undesirable properties, one of them being that for all  $S, \nu$ , and  $\Gamma$ ,

$$\text{if } S, \nu \models_T B\alpha \vee B\beta, \text{ then } S, \nu \models_T B\alpha \wedge B\beta!$$

This follows directly from the peculiar semantics of disjunctions and the fact that, if  $S$  supports  $B\alpha$ , so does any  $S \subseteq \mathcal{S}$ . One of the consequences is that there are no  $S$  and  $\Gamma$  that support the truth of  $\exists x B\alpha \vee \neg \exists x B\alpha$ , a sentence that should come out valid given a suitable semantics.

Intuitively, the problem seems to be the following. If we interpret  $S$  as the actual situation and  $\Gamma$  as the set of situations that are believed possible, then the semantics of  $\vee$  leads us to assume that the truth of any sentence  $\alpha$  containing a disjunction is somehow dependent on the structure of  $S$ . This seems rather counterintuitive in case  $\alpha$  talks about beliefs, since the truth about beliefs (in contrast to knowledge) does not depend on what the world is really like. And certain sentences like  $B\alpha \vee \neg B\alpha$  should come out true no matter what the world and the belief set look like.

We feel that these problems can be overcome if we go back to the more intuitive notions of belief logics, where we have a single situation representing the “actual” situation and a set of situations, which are meant to be those the agent thinks possible. The ideas of t-entailment and, in particular, the notion of global existential quantification should be restricted to this belief set, i.e., only when interpreting a sentence starting with a  $B$ .

In more concrete terms, our aim is a semantics that has as little effect as possible on the standard way of interpreting modal logics in order to retain most of their intuitive appeal. The interpretation of  $B$  should look like:

$$\dots S, s, \dots \models_{\tau} B\alpha \iff \left( \text{condition enabling global exist. quant.} \right) \text{ and } \forall s' \in S \dots S, s' \dots \models_{\tau} \alpha$$

The desire to achieve global existential quantification through a condition that is *local* to the interpretation of the  $B$  operator is guided by the intuition that whatever is done to limit the set of explicit beliefs of the system should have only a minimal effect on the interpretation of logical connectives other than the belief operators. In our case it will result only in a slight twist to the interpretation of  $\exists$ . This will avoid the problems encountered with the non-solution discussed at the beginning of this section.

Before going into the details about how global existential quantification is realized, it should be noted that the sets of situations as used in t-entailment are too restrictive if viewed as belief sets. All members of the set have the same coreference relation, which destroys the difference between non-rigid and rigid designators, a key feature in distinguishing “knowing what” from “knowing that”. For example, if *brother(john)* corefers with the standard name *jim* in all possible situations, then one should conclude that the system knows who he is. Therefore, we allow members of belief sets to have different coreference relations from now on.

In order to get an idea of how to achieve global existential quantification on those sets, assume we are given a situation  $s$  (the “actual” situation), a belief set  $S$ , and a variable map  $\nu$ . Interpreting  $S, s, \nu \models_{\tau} B(\exists x Px \vee \exists y Qy)$  should have the effect of picking two individuals  $a$  and  $b$  global to the entire belief set and replacing the expression by

$$\forall s' \in S \ S, s', \nu_a^x \models_{\tau} Px \text{ or } S, s', \nu_b^y \models_{\tau} Qy \ (*)$$

It turns out that these global choices can be realized via a function  $C$  (called the *choice function*) that takes as arguments the variable map, which takes care of dependencies on leading universal

quantifiers, and the variable name. (\*) can then be replaced by

$$\exists C \forall s' \in S \ S, C, s', \nu_{C(\nu, x)}^x \models_{\mathcal{T}} Px \text{ or } S, C, s', \nu_{C(\nu, y)}^x \models_{\mathcal{T}} Qy,$$

where  $C(\nu, x) = a$  and  $C(\nu, y) = b$ . (Note that the original expression should also carry a choice function for the interpretation of  $\exists$  outside the belief context.)

Therefore, the condition enabling global existential quantification simply amounts to the selection of a choice function before quantifying over the belief set.

So far we have not mentioned what kind of values choice functions range over. From our discussion on quantifying in, it should be clear that the range cannot be standard names, since for a sentence like  $B\exists xPx$  to be true  $Pn$  would have to be true in every element of the belief set, where  $n$  is some standard name. Then, of course, we can choose exactly the same  $n$  to make  $\exists xBPx$  true. The idea is to give a substitutional account of global existential quantification, allowing also non-rigid terms to be substituted. Under that view,  $B\exists xTeacher\ x$  can be true if *brother(john)* is a teacher in all situations, even though this name may refer to different individuals. In a sense, we are saying that the system believes the existence of a teacher if it can *name* one.

Note that, outside the belief context, using terms instead of parameters is redundant, since there the interpretation is done with respect to a *single* situation, and every term corefers with a parameter. This property also allows us to represent “knowing who” properly. All we have to do is to make sure that variables that are used both to talk about the actual situation  $s$  and the beliefs  $S$  as in  $\exists x(Teacher\ x \wedge BTeacher\ x)$  have as a value a standard name (provided by  $s$ ), when a belief is interpreted. Technically, this can be done locally when encountering a  $B$  by converting the variable map  $\nu$  that is active at that moment into a map  $\nu_s$  whose values are just the coreferring standard names (from  $s$ ) of the values of  $\nu$ .

Before presenting the formal semantics, we summarize the major technical changes necessary that allow the transition from t-entailment to explicit beliefs.

- a) The belief set consists of arbitrary first-order situations.
- b) The range of variable maps is extended from  $N$  to  $\mathcal{T}$  to allow variables to be replaced by arbitrary closed terms.
- c) A choice function  $C$  is introduced, mapping pairs of variable maps and variable names into the set of closed terms:

$$C : \mathcal{V} \times V \longrightarrow \mathcal{T}.$$

The set of all choice functions is called  $\mathcal{C}$ .

- d) The conversion of all values of variables into standard names (in the context of explicit or implicit belief). Let  $s$  be a situation,  $\nu$  a variable map.  $\nu_s$  is the variable map s.t.  $\nu_s(x) = n$ , where  $n \in N$  and  $n \equiv_s \nu(x)$ .

Finally, since implicit beliefs are to be interpreted as in a regular possible-worlds semantics, we need the concept of possible worlds compatible with a situation (a straightforward adaptation from [Leve84b]).

$$\mathcal{W}(s) = \{s' \in \mathcal{S} \mid \text{for every k-tuple } \vec{n} \in N^k \text{ and every k-ary predicate P}$$

- a)  $\vec{n}$  is in exactly one of  $T^{s'}(P)$  and  $F^{s'}(P)$
- b) if  $\vec{n} \in T^s(P)$  then  $\vec{n} \in T^{s'}(P)$
- c) if  $\vec{n} \in F^s(P)$  then  $\vec{n} \in F^{s'}(P)$
- d)  $\equiv_{s'} = \equiv_s$  }

$\mathcal{W}(S)$ , the set of all possible worlds compatible with  $S$ , is then the union of all  $\mathcal{W}(s)$  for  $s \in S$ . We can now give the full semantics for the language  $\mathcal{L}$ :

Let  $s \in \mathcal{S}$ ,  $S \subseteq \mathcal{S}$ ,  $\nu \in \mathcal{V}$ , then

1.  $S, s, \nu \models_{\mathcal{T}} \alpha \iff \exists C \in \mathcal{C} \ S, C, s, \nu \models_{\mathcal{T}} \alpha'^6$   
 $S, s, \nu \models_{\mathcal{F}} \alpha \iff \exists C \in \mathcal{C} \ S, C, s, \nu \models_{\mathcal{F}} \alpha'$
2.  $S, C, s, \nu \models_{\mathcal{T}} P\vec{t} \iff \nu(\vec{t}) \in T^s(P)$   
 $S, C, s, \nu \models_{\mathcal{F}} P\vec{t} \iff \nu(\vec{t}) \in F^s(P)$
3.  $S, C, s, \nu \models_{\mathcal{T}} \neg\alpha \iff S, C, s, \nu \models_{\mathcal{F}} \alpha$   
 $S, C, s, \nu \models_{\mathcal{F}} \neg\alpha \iff S, C, s, \nu \models_{\mathcal{T}} \alpha$
4.  $S, C, s, \nu \models_{\mathcal{T}} \alpha \vee \beta \iff S, C, s, \nu \models_{\mathcal{T}} \alpha \text{ or } S, C, s, \nu \models_{\mathcal{T}} \beta$   
 $S, C, s, \nu \models_{\mathcal{F}} \alpha \vee \beta \iff S, C, s, \nu \models_{\mathcal{F}} \alpha \text{ and } S, C, s, \nu \models_{\mathcal{F}} \beta$
5.  $S, C, s, \nu \models_{\mathcal{T}} \exists x\alpha \iff S, C, s, \nu_{\mathcal{C}(s,x)}^x \models_{\mathcal{T}} \alpha$   
 $S, C, s, \nu \models_{\mathcal{F}} \exists x\alpha \iff \forall t \in \mathcal{T} \ S, C, s, \nu_t^x \models_{\mathcal{F}} \alpha$
6.  $S, C, s, \nu \models_{\mathcal{T}} B\alpha \iff \exists C' \in \mathcal{C} \ \forall s' \in S \ S, C', s', \nu_s \models_{\mathcal{T}} \alpha$   
 $S, C, s, \nu \models_{\mathcal{F}} B\alpha \iff S, C, s, \nu \not\models_{\mathcal{T}} B\alpha$
7.  $S, C, s, \nu \models_{\mathcal{T}} L\alpha \iff \forall s' \in \mathcal{W}(S) \ \exists C' \in \mathcal{C} \ S, C', s', \nu_s \models_{\mathcal{T}} \alpha$   
 $S, C, s, \nu \models_{\mathcal{F}} L\alpha \iff S, C, s, \nu \not\models_{\mathcal{T}} L\alpha$

In the remainder of this section, we discuss how this semantics relates to classical Tarskian semantics and look at some of the properties of explicit and implicit belief. We start by defining *validity* for a formula  $\alpha \in \mathcal{L}$  as

$$\models \alpha \iff \text{for all } S, s \in \mathcal{W}(S), \nu \in \mathcal{V} \ S, s, \nu \models_{\mathcal{T}} \alpha$$

<sup>6</sup> $\alpha'$  is the same as  $\alpha$  except that all variable names are distinct. This condition is needed, since the choice functions depend on variable names. Note that in subsequent proofs we ignore the renaming step in order to preserve clarity.

Given that the semantics uses a non-standard way of interpreting  $\exists$ , one should feel somewhat uncomfortable with this definition of validity. Namely, does it have anything to do with what we normally call a valid sentence? The first task will therefore be to show that this notion of validity does in fact reduce to standard Tarskian validity if a sentence contains no modal operator. Furthermore, we will show that the same holds for truths about beliefs, something we found missing in our first extension to t-entailment.

Given a Tarskian world  $w = \langle T^w, F^w, \equiv_w \rangle$  ( $\in \mathcal{W}(S)$ ), we define a Tarskian style interpretation function  $I_w : \mathcal{L}_1 \rightarrow \{T, F\}$  as follows:

$$I_w(P\vec{n}) = T \iff \vec{n} \in T^w(P)$$

$$I_w(\neg\alpha) = T \iff I_w(\alpha) = F$$

$$I_w(\alpha \vee \beta) = T \iff I_w(\alpha) = T \text{ or } I_w(\beta) = T$$

$$I_w(\exists x\alpha) = T \iff \text{for some } n \in N \ I_w(\alpha^{x/n}) = T$$

$$I_w(\alpha^{x/t_1}) = I_w(\alpha^{x/t_2}), \text{ where } t_1 \equiv_w t_2 \text{ and } x \text{ is the only free variable in } \alpha.$$

Tarskian validity for a sentence  $\alpha \in \mathcal{L}_1$  is then defined as

$$\models^T \alpha \text{ if and only if for all } w \in \mathcal{W}(S) \ I_w(\alpha) = T.$$

The following lemma states a strong correspondence between our semantics and Tarskian semantics.

**Lemma 1** *For any sentence  $\alpha \in \mathcal{L}_1, w \in \mathcal{W}(S), S$ , and  $\nu$*

$$\begin{aligned} I_w(\alpha) = T &\iff S, w, \nu \models_T \alpha \\ I_w(\alpha) = F &\iff S, w, \nu \models_F \alpha \end{aligned}$$

**Proof :** By induction on  $\alpha$ . Since both model theories coincide in their definitions for atoms, negations, and disjunctions, the only interesting case is existential quantification. We show the only if direction for both statements, where  $\alpha$  starts with a  $\exists$ .

Let  $I_w(\exists x\alpha) = T$ , then  $I_w(\alpha^{x/n}) = T$  for some  $n \in N$ , implying  $\exists C \ S, C, w, \nu \models_T \alpha^{x/n}$ . From lemma 9 in the appendix we get  $\exists C' \ S, C', w, \nu_n^x \models_T \beta$ , but then  $\exists C'' \ S, C'', w, \nu_{C''(\nu, x)}^x \models_T \beta$  (choose  $C'' = C'$  except  $C''(\nu, x) = n$ ) implying  $\exists C'' \ S, C'', w, \nu \models_T \exists x\alpha$ , from which  $S, s, \nu \models_T \exists x\alpha$  follows.

Now let  $I_w(\exists x\alpha) = F$ , then  $\forall n \in N \ I_w(\alpha^{x/n}) = F$ , which implies  $\forall n \in N \ \exists C \ S, C, s, \nu \models_F \alpha^{x/n}$ . From lemma 9 it follows that  $\forall n \in N \ \exists C' \ S, C', s, \nu_n^x \models_F \alpha$  and by lemma 8 of the appendix,  $\forall t \in \mathcal{T} \ \exists C'' \ S, C'', s, \nu_t^x \models_F \alpha$ . Finally, with lemma 10 of the app. we obtain  $\exists C^* \ \forall t \in \mathcal{T} \ S, C^*, s, \nu_t^x \models_F \alpha$ , which implies  $S, s, \nu \models_F \exists x\alpha$ . ■

From the previous lemma it follows immediately that the two notions of validity coincide for sentences not containing modal operators.

**Theorem 2** *For any  $\alpha \in \mathcal{L}_1, \models^T \alpha \iff \models \alpha$*

From theorem 2 we derive immediately that, for example,  $(Pa \vee Pb) \supset \exists xPx$  is valid in our logic, that is, existential quantification outside modal operators behaves normally.

For sentences that do contain modal operators, we proceed as follows. A formula  $\alpha'$  is called a *first-order substitution* of  $\alpha$ , if it is obtained from  $\alpha$  by substituting all occurrences of  $B\beta$  ( $L\beta$ ) in  $\alpha$  by a new predicate  $P_{B\beta}\vec{x}$  ( $P_{L\beta}\vec{x}$ ), where  $\vec{x}$  is the vector of all free variables in  $\beta$ .

**Lemma 3** For any  $\alpha \in \mathcal{L}$ , its first-order substitution  $\alpha'$ , and for any  $S, C, s \in \mathcal{W}(S)$ , and  $\nu$

$$\begin{aligned} S, C, s, \nu \models_{\tau} \alpha &\iff S, C, s', \nu \models_{\tau} \alpha' \\ S, C, s, \nu \models_{\text{F}} \alpha &\iff S, C, s', \nu \models_{\text{F}} \alpha', \end{aligned}$$

where  $s = s'$  except for all  $P_{B\beta}$  in  $\alpha'$  we define (and similarly for  $P_{L\beta}$ )

$$\begin{aligned} T^{s'}(P_{B\beta}) &= \{\vec{n} \mid S, C, s, \nu_{\vec{n}} \models_{\tau} B\beta\} \\ F^{s'}(P_{B\beta}) &= \{\vec{n} \mid S, C, s, \nu_{\vec{n}} \models_{\text{F}} B\beta\} \end{aligned}$$

**Proof :** Trivial. Note that  $s'$  itself is a possible world, since exactly one of  $S, C, s, \nu_{\vec{n}} \models_{\tau} B\beta$  or  $S, C, s, \nu_{\vec{n}} \models_{\text{F}} B\beta$  is true for all  $\vec{n}$ . ■

**Theorem 4** Let  $\alpha'$  be a first-order substitution of a sentence  $\alpha \in \mathcal{L}$ .

If  $\models^{\tau} \alpha'$ , then  $\models \alpha$ .

**Proof :** Assume  $\models^{\tau} \alpha'$  and  $\not\models \alpha$ . Then there are  $S, s$ , and  $\nu$  s.t. for all  $C$ ,  $S, C, s, \nu \not\models_{\tau} \alpha$ , but then  $S, C, s', \nu \not\models_{\tau} \alpha'$ . From lemma 1 it follows that  $I_{s'}(\alpha') \neq T$ , and therefore  $\not\models^{\tau} \alpha'$  contradicting our assumption. ■

As a corollary, for example, we get immediately that  $\forall xB\alpha \vee \neg\forall xB\alpha$  is valid. Overall, then,  $\mathcal{L}$  is "upward compatible" with classical first-order logic.

We can now turn to results concerning implicit and explicit beliefs and, of course, quantifying in. First of all, implicit beliefs are modelled according to possible worlds semantics. In particular, the operator  $L$  behaves much like the  $K$  operator in Levesque's logic  $K\mathcal{L}$  (without equality)[Leve82]. Logical omniscience of implicit beliefs is captured in the following theorem.

**Theorem 5**

- |  |  |
|--|--|
| a) If $\models \alpha$ then $\models L\alpha$ for $\alpha \in \mathcal{L}_1$ | <i>Implicit belief of first-order valid formulas</i> |
| b) $\models L\alpha \wedge L(\alpha \supset \beta) \supset L\beta$           | <i>Modus Ponens</i>                                  |
| c) $\models \forall xL\alpha \supset L\forall x\alpha$                       | <i>Universal Generalization</i>                      |

**Proof :** We prove only (a) here. If  $\models \alpha$ , then for all  $S, s \in \mathcal{W}(S)$ , and  $\nu$  there is a  $C$  s.t.  $S, C, s, \nu \models_{\tau} \alpha$ . Since  $\mathcal{W}(S) \subseteq \mathcal{W}(S)$ , we get immediately  $\forall s' \in \mathcal{W}(S) \exists C$  s.t.  $S, C, s', \nu \models_{\tau} \alpha$ . With  $\nu_s(x) \equiv_s \nu(x)$  the result follows from lemma 8 (appendix).

The idea to keep in mind is that in any interpretation of  $L$  we can choose a different choice function for every possible world, which reduces to a Tarskian interpretation as we have seen in theorem 2. ■

Explicit and implicit beliefs have the proper connection, i.e., every explicit belief is also implicitly believed.

**Theorem 6**  $\models B\alpha \supset L\alpha$

**Proof** : It is easy to see that whenever a situation  $s$  supports the truth (falsity) of  $\alpha$ , so does every possible world compatible with  $s$  (even with the same choice function). Also, having a choice function ranging over all elements of the belief set (as for  $B\alpha$ ) is a much stronger requirement than being able to choose a different  $C$  for every element of the set (as for  $L\alpha$ ). ■

Finally, we have the following properties for explicit and implicit beliefs with respect to quantifying in:

**Theorem 7** For any  $\alpha$  in  $\mathcal{L}_1$  and any parameter  $n$  and closed term  $a$

- a)  $\models \exists x B\alpha \supset B\exists x\alpha$  and  $\models \exists x L\alpha \supset L\exists x\alpha$
- b)  $\models B\alpha^{x/a} \supset B\exists x\alpha$  and  $\models L\alpha^{x/a} \supset L\exists x\alpha$
- c)  $\models B\alpha^{x/n} \supset \exists x B\alpha$  and  $\models L\alpha^{x/n} \supset \exists x L\alpha$
- d)  $\models B\forall x\alpha \supset \forall x B\alpha$  and  $\models L\forall x\alpha \supset \forall x L\alpha$

**Proof** : We only look at explicit beliefs here.

a)  $\models \exists x B\alpha \supset B\exists x\alpha$

It suffices to show that, if  $S, C, s, \nu \models_{\tau} \exists x B\alpha$ , then  $S, C, s, \nu \models_{\tau} B\exists x\alpha$  for any  $S, C, s$ , and  $\nu$ . Let  $S, C, s, \nu \models_{\tau} \exists x B\alpha$ , then  $S, C, s, \nu_{C(\nu, x)}^x \models_{\tau} B\alpha$  and therefore  $\exists C'$  s.t.  $\forall s' \in S$

$S, C', s', (\nu_{C'(\nu, x)}^x)_s \models_{\tau} \alpha$ . Let  $C(\nu, x) \equiv_s n$  for some  $n \in N$ , then define  $C''$  as  $C'$  except  $C''(\nu, x) = n$ , from which the result follows.

b)  $\models B\alpha^{x/a} \supset B\exists x\alpha$

As before, let  $S, C, s, \nu \models_{\tau} B\alpha^{x/a}$ , then  $\exists C'$  s.t.  $\forall s' \in S$   $S, C', s', \nu_s \models_{\tau} \alpha^{x/a}$ , and therefore (by lemma 9 of the appendix)  $\exists C''$  s.t.  $\forall s' \in S$   $S, C'', s, (\nu_s)_a^x \models_{\tau} \alpha$ . Define  $C^*$  as  $C''$  except that  $C^*(\nu_s, x) = a$ , from which the result follows.

c)  $\models B\alpha^{x/n} \supset \exists x B\alpha$

It suffices to show that for all  $C$  s.t.  $C(\nu, x) = n$ , if  $S, C, s, \nu \models_{\tau} B\alpha^{x/n}$ , then  $S, C, s, \nu \models_{\tau} \exists x B\alpha$ . Let  $S, C, s, \nu \models_{\tau} B\alpha^{x/n}$ , then  $\exists C' \forall s' \in S$   $S, C', s', \nu_s \models_{\tau} \alpha^{x/n}$ , and therefore (by lemma 9)  $\exists C'' \forall s' \in S$   $S, C'', s', (\nu_s)_n^x \models_{\tau} \alpha$ . Since  $(\nu_s)_n^x = (\nu_n^x)_s$ , we get  $S, C, s, \nu_n^x \models_{\tau} B\alpha$ . Then, since we assumed that  $C(\nu, x) = n$ , it follows that  $S, C, s, \nu \models_{\tau} \exists x B\alpha$ .

d)  $\models B\forall x\alpha \supset \forall x B\alpha$

Again, the proof is not hard. On the left hand side, the quantifier ranges over all closed terms, whereas the right hand side essentially restricts the quantification to parameters. ■

In addition, it is easy to show that "knowing that" is strictly weaker than "knowing what" for both implicit and explicit beliefs: There are  $\alpha \in \mathcal{L}_1$  and non-rigid terms  $a$  s.t.

- a)  $\not\models B\exists x\alpha \supset \exists x B\alpha$  and  $\not\models L\exists x\alpha \supset \exists x L\alpha$
- b)  $\not\models B\alpha^{x/a} \supset \exists x B\alpha$  and  $\not\models L\alpha^{x/a} \supset \exists x L\alpha$

As an example for explicit belief, let  $\alpha = Px$  and  $S = \{s_1, s_2\}$ , where  $s_1 \models_{\tau} P1$  and  $s_2 \models_{\tau} P2$ , and  $a \equiv_{s_1} 1$  and  $a \equiv_{s_2} 2$  for some constant  $a$ . In addition, let no other term corefer with 1 or 2 in either situation.

The limited reasoning power on explicit beliefs can be summarized as follows: There are  $\alpha$  and  $\beta \in \mathcal{L}$  such that

$$\text{a) } \not\models B\alpha \wedge B(\alpha \supset \beta) \supset B\beta$$

No Modus Ponens

$$\text{b) } \not\models \forall x B\alpha \supset B\forall x\alpha$$

No Universal Generalization

The failure of modus ponens follows directly from the fact that the propositional subset of our logic reduces to Levesque's logic in [Leve84b].

An example why universal generalization fails is the following: Let  $\alpha = \exists y Pxy$  and  $S = \{s_1, s_2\}$ . Both  $s_1$  and  $s_2$  support the truth of  $P12, P23, P34, \dots$ , and for some constant  $a$ ,  $a \equiv_{s_1} 1$  and  $a \equiv_{s_2} 2$ . All other terms corefer with parameter 1 in both situations. Then certainly  $S$  supports  $\forall x B\alpha$ , since only parameters are relevant here. However, there is no term we could substitute for  $y$  s.t.  $S$  supports  $B\exists y\alpha^{x/a}$ .

As a concluding remark we note that the notion of explicit belief we have put forward in this paper allows an agent to hold beliefs with varying degrees of strength.

A very weak kind is the belief of a disjunction, since, for example,  $B(Pa \vee Pb) \supset B\exists x Px$  is not valid. (As in t-entailment, it is easy to construct a belief set, where this sentence fails. Simply let  $S = \{s_1, s_2\}$ , where  $s_1$  supports only the truth of  $Pn$  and  $s_2$  the truth of  $Pm$  for distinct parameters  $n$  and  $m$ .) On the other hand, note that  $B(Pa \vee Pb) \supset L\exists x Px$  is valid.

Although explicit beliefs of type "knowing that" are stronger than disjunctive beliefs, they are strictly weaker than those of type "knowing what":

$$\not\models B\exists x\alpha \supset \exists x B\alpha \quad \text{vs.} \quad \models \exists x B\alpha \supset B\exists x\alpha$$

Finally, a belief about the existence of someone having some property  $\alpha$  is strongest, if the system can pinpoint a particular individual with that property:  $\models B\alpha^{x/n} \supset \exists x B\alpha$ , where  $n$  is a parameter.

## Conclusion

A major focus of this paper has been the problem of quantifying in as it relates to logics of explicit and implicit belief. We have presented a semantics for explicit belief that preserves the ability to distinguish between "knowing what" and "knowing that", but is otherwise much weaker than implicit belief.

Work on the semantics and its properties is still in progress. Recently, we have obtained a decidability result that closely mirrors that of Patel-Schneider's t-entailment for formulas of the form  $B\alpha \supset B\beta$ , where  $\alpha$  and  $\beta$  are in prenex normal form (that is, all quantifiers are moved to the front). We hope to report on these issues and present a complete axiomatization in a later paper.

An interesting phenomenon of our semantics that needs further study is the fact that not all logically equivalent forms of a formula are believed explicitly. For example,  $B(\forall x \exists y (Px \vee Qy)) \supset B(\exists y \forall x (Px \vee Qy))$  is not valid.

Finally, two further areas of investigation are nested beliefs and a belief logic with equality. For example, it seems reasonable to expect  $\models B(n = n)$  to hold for any parameter  $n$ . On the other hand, it is not clear whether  $\models B(t_1 = n) \wedge B(t_2 = n) \supset B(t_1 = t_2)$  should come out true for arbitrary terms  $t_1$  and  $t_2$ .

### Acknowledgements

I am indebted to Hector Levesque for many enlightening discussions on this topic and his comments on earlier drafts of this paper. I would also like to thank Alan Frisch, Jim des Rivières, Murray Mazer, Greg McArthur, and Martin Stanley. Their comments have led to numerous improvements of the paper. Financial support was gratefully received from the World University Service of Canada.

### References

- [AnBe75] Anderson, A. R. and Belnap, N. D., *Entailment, The Logic of Relevance and Necessity*, Princeton University Press, 1975.
- [Beln77] Belnap, N. D., A Useful Four-Valued Logic, in G. Epstein and J. M. Dunn (eds.), *Modern Uses of Multiple-Valued Logic*, Reidel, 1977.
- [Eber74] Eberle, R. A., A Logic of Believing, Knowing and Inferring, *Synthese* 26, 1974, pp. 356-382.
- [FaHa85] Fagin, R. and Halpern, J. Y., Belief, Awareness, and Limited Reasoning: Preliminary Report, *Proc. Int. Joint Conf. on AI*, August 1985, pp. 491-501.
- [Freg92] Frege, G., Über Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik*, Vol. 100, 1892, pp. 25-50.
- [Fris85] Frisch, A. M., Using Model Theory to Specify AI Programs, *Proc. Int. Joint Conf. on AI*, August 1985, pp. 148-154.
- [Hint62] Hintikka, J., *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, Cornell University Press, 1962.
- [Hint71] Hintikka, J., Semantics for Propositional Attitudes, in L. Linsky (ed.), *Reference and Modality*, Oxford University Press, Oxford, 1971.
- [HuCr68] Hughes, G. E. and Cresswell, M. J., *An Introduction to Modal Logic*, Methuen and Company Ltd., London, England, 1968.
- [Kono84] Konolige, K., Belief and Incompleteness, SRI Artificial Intelligence Note 319, SRI International, Menlo Park, 1984.
- [Leve82] Levesque, H. J., A Formal Treatment of Incomplete Knowledge Bases, Tech. Report No. 3, Fairchild Lab. for AI Research, Palo Alto, 1982.
- [Leve84a] Levesque, H. J., Foundations of a Functional Approach to Knowledge Representation, *Artificial Intelligence*, Vol. 23, 1984, pp. 155-212.
- [Leve84b] Levesque, H. J., A Logic of Implicit and Explicit Belief, Tech. Rep. No. 32, Fairchild Lab. for AI Research, Palo Alto, 1984.
- [Leve] Levesque, H. J., Global and Local Consistency and Completeness of Belief, in preparation.
- [Lins71] Linsky, L. (ed.), *Reference and Modality*, Oxford University Press, Oxford, 1971.

- [Moor80] Moore, R. C., Reasoning about Knowledge and Action, Technical Note 181, SRI International, Menlo Park, 1980.
- [MoHe79] Moore, R. C. and Hendrix, G., Computational Models of Beliefs and the Semantics of Belief-Sentences, Technical Note 187, SRI International, Menlo Park, 1979.
- [Pate85] Patel-Schneider, P. F., A Decidable First-Order Logic for Knowledge Representation, *Proc. Int. Joint Conf. on AI*, August 1985, pp. 455-458, (an extended version is in preparation as: AI Tech. Report No. 45, Schlumberger Palo Alto Research).

## Appendix

We state three technical lemmas about choice functions, which we used in previous proofs. Loosely speaking, these lemmas express that choice functions are very flexible with respect to variable maps. The first lemma says that choice functions are basically equivalent modulo the substitution of coreferring terms in a variable map.

**Lemma 8** *Given a situation  $s$ , a variable  $x$ , and a closed term  $u$ , define choice functions  $C_1$  and  $C_2$  s.t. for all  $v \in \mathcal{T}$ , where  $u \equiv_s v$ ,  $C_1(\nu_u^x, y) = C_2(\nu_v^x, y)$  for all  $\nu$  and  $y$ . Then, for all  $S, \nu, v$  with  $u \equiv_s v$  and  $\alpha$ :*

$$\begin{aligned} S, C_1, s, \nu_u^x \models_{\mathcal{T}} \alpha &\iff S, C_2, s, \nu_v^x \models_{\mathcal{T}} \alpha \\ S, C_1, s, \nu_u^x \models_{\mathcal{F}} \alpha &\iff S, C_2, s, \nu_v^x \models_{\mathcal{F}} \alpha \end{aligned}$$

Lemma 9 tells us that using a variable map is equivalent to direct substitution of values into a formula, provided no substitutions occur within beliefs.

**Lemma 9** *For closed terms  $u$  and  $v$  define two choice functions  $C_u$  and  $C_v$  s.t.  $C_u(\nu_u^x, y) = C_v(\nu_v^x, y)$  for all  $y$  and  $\nu$  and some  $x$ . Then, for any  $\alpha \in \mathcal{L}$  with  $x$  free in  $\alpha$ , but not in the scope of a  $B$  or  $L$ , and any  $S, s$  and  $\nu$  with  $\nu(x) = v$ :*

$$\begin{aligned} S, C_u, s, \nu_u^x \models_{\mathcal{T}} \alpha &\iff S, C_v, s, \nu \models_{\mathcal{T}} \alpha^{x/u} \\ S, C_u, s, \nu_u^x \models_{\mathcal{F}} \alpha &\iff S, C_v, s, \nu \models_{\mathcal{F}} \alpha^{x/u} \end{aligned}$$

The next lemma shows an important case of merging a set of choice functions with respect to a partitioning of variable maps over one variable.

**Lemma 10** *For any closed term  $t$  let  $C_t$  be some choice function. Then, for any  $x$  and  $\alpha$  there is a  $C$  s.t. for all  $u \in \mathcal{T}$ :*

$$\begin{aligned} S, C_u, s, \nu_u^x \models_{\mathcal{T}} \alpha &\iff S, C, s, \nu_u^x \models_{\mathcal{T}} \alpha \\ S, C_u, s, \nu_u^x \models_{\mathcal{F}} \alpha &\iff S, C, s, \nu_u^x \models_{\mathcal{F}} \alpha \end{aligned}$$

**Proof :** Let  $C(\nu, y) = C_t(\nu, y)$  for  $\nu(x) = t$ . The rest follows by a simple induction argument on  $|\alpha|$ . ■