An Approach to Modeling Indexicality in Action and Communication*

Yves Lespérance
School of Computing Science
Simon Fraser University
Burnaby, B.C. V5A 1S6 Canada
yves@cs.sfu.ca

Abstract

While many uses of context sensitive expressions in language are just communicative shortcuts, ways of succinctly referring to the relevant entities, there are also cases where the information that needs to be communicated is intrinsically indexical. To model this, one needs a formal account of communication that relates the context sensitivity of language to that of mental states and action. In this paper, we develop a preliminary version of such an account, using our theory of indexical knowledge and action as a foundation. The account includes a formalization of the notion of mutual indexical knowledge, a specification of the semantics of a communication language that includes indexicals, and a formalization of the pragmatics of declarative utterances that supports reasoning about the ability of agents to achieve goals by making such utterances.

Introduction

In previous work, we developed a formal account of the relationship between indexical (context sensitive) knowledge and non-linguistic action [12, 10]. We are now investigating the relationship between the indexicality involved in linguistic communication and that implicated in thought and non-linguistic action. In [12, 10], we argued that the knowledge required for many types of non-linguistic action (e.g. robot navigation, manipulation) is indexical. This leads us to observe that while many uses of indexical expressions in language are only communicative shortcuts, ways of succinctly referring to the relevant entities, there are also cases where the information that needs to be communicated is intrinsically indexical. For example, I can help you get to my place by telling you where it is *relative to your current position*. Such situations, where agents cooperatively accomplish a task using both linguistic and physical acts, are good domains for studying the interaction between the indexicality involved in language and that involved thought and non-linguistic action.

In this paper, we describe a preliminary version of a theory that can model this kind of situation. The theory accounts for the fact that neither the speaker nor the hearer need know what objective proposition was expressed by an utterance. It also accounts for the perspective shift that occurs in comprehension — how speaker-relative knowledge gets mapped into hearer-relative knowledge. A key feature of the proposed approach is that it tries to specify the cognitive significance of the numerous indexicals in natural language in terms of a few indexical knowledge primitives, that is, the primitives **self** and **now** included in our earlier work. We think that the theory also shows that the "direct reference" account of the semantics of indexicals is compatible with the view that their cognitive significance involves descriptive concepts. We feel that dealing with such semantics issues within an integrated account of semantics and pragmatics can bring new perspectives onto the associated problems.

After outlining the formalism in which our theory is specified in the next section, we will in turn present a formalization of the notion of mutual indexical knowledge, a semantic account of a communication language that includes indexicals, and a simple formalization of the pragmatics of declaratives that accounts for the ability of agents to achieve goals through such utterances.

Overview of the Formalism

Our theory is specified in a many-sorted quantified modal logic that is an extended version of the formalism used in our earlier work [12, 10]. The main difference between our formalism and more standard ones is that it allows attributions of indexical knowledge to be expressed, for example, that Rob knows that he himself was holding a cup five minutes ago.[1] In such cases, what is known is a "proposition" that is relative. It may be relative to the knower, or to the time of the knowing, or

*This research was made possible in part by financial support from the Institute for Robotics and Intelligent Systems and NTT Corporation. Some of the work was carried out at NTT Basic Research Laboratories, Musashino, Japan.

[1] As well, this can be distinguished from having objective knowledge, such as Rob's knowing that he was holding a cup at some specified time, say, 4:37pm.

perhaps to other aspects of the context. To handle this, our language includes two special terms: **self**, which denotes the current agent, and **now**, which denotes the current time; we call these terms *primitive indexicals*. Note that **self** and **now** are *not* intended to be formal counterparts of similarly sounding English words and often behave quite differently from any such words. Non-logical (domain-dependent) symbols may also depend on the current agent and time for their interpretation, for example, $\exists x \textsc{Holding}(x)$ may express the fact that the current agent is currently holding something — we say that such symbols are *non-primitive indexicals*. Our semantics handles this by interpreting terms and formulas with respect to *indices*, which consist of a possible-world (modeling the objective circumstances), an agent and a time (modeling the context).

The language includes terms of several different sorts. There are terms for ordinary individuals (as usual — sort i), agent terms (sort a), temporal terms (sort t), and action terms (sort d). For each sort, there are both variables and function symbols (i.e., functions whose values are of the proper sort); as usual, constants are taken to be 0-ary function symbols. There are also terms that stand for expressions in the language used by agents to communicate. Since the communication language is largely a subset of the logic itself, it will be introduced after the rest of the logic.

The atomic formulas include predications using predicate symbols and terms, written $\boldsymbol{R}(\theta^i_1, \ldots \theta^i_n)$, which are used to assert that $\theta^i_1, \ldots \theta^i_n$ stand in static relation $\boldsymbol{R}$ at the current time for the current agent (θ^σ stands for a term of sort σ). We also have equality expressions $(\theta_1 = \theta_2)$, between terms of the same sort, as well as expressions of temporal precedence $(\theta^t_1 < \theta^t_2)$. We assume that time is linearly ordered. Finally, $\mathbf{Does}(\theta^d, \theta^t)$ is used to assert that the current agent does action θ^d starting from the current time and ending at time θ^t.

Non-atomic formulas may be composed using the standard boolean connectives and quantifiers as well as a set of modal operators. $\mathbf{At}(\theta^t, \varphi)$ means that φ holds at time θ^t, that is, when θ^t is taken to be the current time. $\mathbf{By}(\theta^a, \varphi)$ means that φ holds when θ^a is taken to be the current agent. $\Box\varphi$ means that φ is historically necessary at the current time, that is, that φ now holds in all possible courses of events that are identical to the current one up to the current time. We also introduce a dual to $\Box$: $\Diamond\varphi \stackrel{\text{def}}{=} \neg\Box\neg\varphi$.

$\mathbf{Know}(\varphi)$ is used to represent the fact that the current agent knows at the current time that φ. If φ contains indexical elements, $\mathbf{Know}(\varphi)$ should be taken as attributing indexical knowledge, that is, knowledge the agent has about himself and the current time. For example, $\mathbf{Know}(\textsc{Holding}(x))$ could mean that the agent knows that he himself is currently holding the object denoted by x. The semantics for **Know** is a simple generalization of the standard possible-world scheme [9, 6]. The knowledge accessibility relation K is taken to hold over *indices* rather than plain possible worlds. Informally, $\langle\langle \mathsf{w}, \mathsf{a}, \mathsf{t}\rangle, \langle \mathsf{w}', \mathsf{a}', \mathsf{t}'\rangle\rangle \in \mathsf{K}$ if and only if as far as agent a at time t in world w knows, it may be the case that w' is the way the world actually is *and* he is a' *and* the current time is t'. Thus, we allow an agent to be uncertain not only about what world he is in, but also about who he is and what time it is. We assume that the knowledge accessibility relation K is reflexive and transitive, meaning that **Know** obeys the principles of modal logic S4. We also assume that agents have perfect memory and always know what actions they have done. Note that the more common version of **Know** that specifies which agent knows a proposition is defined as follows: $\mathbf{Know}(\theta^a, \varphi) \stackrel{\text{def}}{=} \mathbf{By}(\theta^a, \mathbf{Know}(\varphi))$; similar definitions can be provided for other operators such as **Does**, **Can**, etc.

To this we add the operator **Goal** drawn from Cohen and Levesque's formalism [3, 4]. $\mathbf{Goal}(\varphi)$ means that φ follows from the agent's goals at the current time. As for **Know**, we interpret **Goal** as a relation over indices, thus handling indexical goals (e.g., an agent's goal that he himself be holding a cup five minutes from now). We assume that the goal accessibility relation G is serial, meaning that **Goal** obeys the principles of modal logic KD. We also assume $\mathsf{G} \subseteq \mathsf{K}$, that is, that the indices that are consistent with what the agent has chosen are not ruled out by his knowledge (realism).

As mentioned earlier, the communication language handled is largely a subset of the logic itself; for example, $\mathtt{At}(\mathtt{now}+5, \mathtt{pos}(\mathtt{I}) = \mathtt{here})$ is a term that stands for that very sentence of the communication language (and $\textsc{uttrDcl}(\mathtt{At}(\mathtt{now}+5, \mathtt{pos}(\mathtt{I}) = \mathtt{here}))$ is an action term that stands for the act of making a declarative utterance of such a sentence). The communication language can be viewed as a kind of logical form for a natural language or as an artificial language for, say, communicating robots. Terms that stand for communication language expressions fall into five sorts: those that stand for communication language individual terms (sort I), those that stand for agent terms (sort A), those that stand for temporal terms (sort T), those that stand for action terms (sort D), and those those that stand for communication language formulas (sort F). `I` and `you` are constants of sort A, `here` is a constant of sort I, and `now` is a constant of sort T; all of these are intended to behave like the corresponding English indexical. There are also function symbols of these sorts (e.g., `fatherOf(paul)`). A denumerably infinite subset of these is taken to stand for communication language variables of the associated sort: $\mathtt{I}_1, \mathtt{I}_2, \ldots$ for individuals, $\mathtt{A}_1, \mathtt{A}_2, \ldots$ for agents, etc. We also include terms of the form $\mathtt{dthat}(\theta)$ (where θ is of sort I, A, T, or D) that behave like a referential use of θ (`dthat` is taken from Kaplan [8, 7]). The sort "communication language formula" (F) includes terms of the form: $\mathtt{R}(\theta^I_1, \ldots, \theta^I_n)$, $\mathtt{Does}(\theta^D, \theta^T)$, $(\theta_1 = \theta_2)$, $(\theta^T_1 < \theta^T_2)$, $\neg\theta^F$, $\theta^F_1 \wedge \theta^F_2$, $\forall \mathtt{V} \theta^F$, $\mathtt{At}(\theta^T, \theta^F)$, $\mathtt{By}(\theta^A, \theta^F)$, and $\Box\theta^F$. These terms are treated as non-logical (domain-dependent) symbols by the logic; their "semantics" will be specified axiomatically in a later section.

To talk more easily about a wider class of actions, it is useful to extend the use of **Does** to a new syntactic category, that of *action expressions*. These include action terms as mentioned earlier, which represent simple actions, **skip**, which represents the empty action and takes no time, $(\delta_1;\delta_2)$, which represents the sequential composition of the actions δ_1 and δ_2, and $\mathbf{if}(\varphi,\delta_1,\delta_2)$, which represents the action that consists in doing action δ_1 if the condition φ holds, and in doing action δ_2 otherwise. Formulas of the form $\mathbf{Does}(\delta,\theta^t)$ where δ is an action expression, can be thought of as abbreviations that reduce to formulas where **Does** ranges only over the simple action terms, in the obvious way. A bounded form of "while loop" is also defined in terms of conditionals and sequences.

Let us also define some dynamic-logic-style operators that will be used in our formalization of ability. $\mathbf{AfterNec}(\delta,\varphi)$, which is intended to mean "φ must hold after δ", is defined inductively as follows:

$\mathbf{AfterNec}(\theta^d,\varphi) \stackrel{\text{def}}{=} \Box\forall v^t(\mathbf{Does}(\theta^d,v^t) \supset \mathbf{At}(v^t,\varphi))$, where v^t is a temporal variable that does not occur free in φ

$\mathbf{AfterNec}(\mathbf{skip},\varphi) \stackrel{\text{def}}{=} \varphi$

$\mathbf{AfterNec}((\delta_1;\delta_2),\varphi) \stackrel{\text{def}}{=}$ $\mathbf{AfterNec}(\delta_1,\mathbf{AfterNec}(\delta_2,\varphi))$

$\mathbf{AfterNec}(\mathbf{if}(\varphi_c,\delta_1,\delta_2),\varphi) \stackrel{\text{def}}{=}$ $(\varphi_c \supset \mathbf{AfterNec}(\delta_1,\varphi)) \wedge (\neg\varphi_c \supset \mathbf{AfterNec}(\delta_2,\varphi))$

Also, let $\mathbf{PhyPoss}(\delta) \stackrel{\text{def}}{=} \neg\mathbf{AfterNec}(\delta,\mathbf{False})$. $\mathbf{PhyPoss}(\delta)$ is intended to mean that it is "physically possible" for **self** to do action δ next (even though he may not be able to do it because he does not know what primitive actions δ stands for). **True** (**False**) stands for some tautology (contradiction).

Our formalization of ability, based on that of Moore's [14], says that the agent is able to achieve the goal φ by doing action δ, formally $\mathbf{Can}(\delta,\varphi)$, if and only if he can do action δ and knows that after doing δ, the goal φ must hold: $\mathbf{Can}(\delta,\varphi) \stackrel{\text{def}}{=} \mathbf{CanDo}(\delta) \wedge \mathbf{Know}(\mathbf{AfterNec}(\delta,\varphi))$. $\mathbf{CanDo}(\delta)$ is defined inductively as follows:[2]

$\mathbf{CanDo}(\theta^d) \stackrel{\text{def}}{=}$ $\exists v^d\,\mathbf{Know}(v^d=\theta^d) \wedge \mathbf{Know}(\mathbf{PhyPoss}(\theta^d))$, where action variable v^d does not occur free in θ^d

$\mathbf{CanDo}(\mathbf{skip}) \stackrel{\text{def}}{=} \mathbf{True}$

$\mathbf{CanDo}(\delta_1;\delta_2) \stackrel{\text{def}}{=} \mathbf{Can}(\delta_1,\mathbf{CanDo}(\delta_2))$

[2] This way of defining **Can** is preferable to the one in [10, 12] as it separates the knowledge prerequisites involving the goal from the rest. The definitions of **AfterNec** and **PhyPoss** given here are also changed; they now behave exactly as their dynamic logic [5] counterparts do. See [11] for further discussion.

$\mathbf{CanDo}(\mathbf{if}(\varphi,\delta_1,\delta_2)) \stackrel{\text{def}}{=} (\mathbf{Know}(\varphi) \wedge \mathbf{CanDo}(\delta_1)) \vee (\mathbf{Know}(\neg\varphi_c) \wedge \mathbf{CanDo}(\delta_2))$

Note that we eliminate Moore's requirement that the agent know who he is; instead, we require indexical knowledge (see [13] for a discussion of why this is better). Also, the fact that our account of ability is based on a more expressive temporal logic allows it to deal with actions whose prerequisites or effects involve knowledge of absolute times and knowing what time it is.

Mutual Indexical Knowledge

Analyses of communication use the notion of mutual knowledge. In the usual definitions of this notion, it is assumed that each agent knows who he is and knows who the other agent(s) involved is (are). This ignores the fact that agents may only know of themselves and the other agent under some indexical description. For example, suppose that John experiences love at first sight for a beautiful stranger and immediately proceeds to say "I love you" to this person. Yet they do not know each other's identity — they haven't been introduced. John only knows of the hearer as "the agent who is attending to me" or some other such indexical description, and similarly, the hearer only knows of him as the "the agent whom I am attending to". Let's define the hearer's notion of the speaker, **SforH**, as "the one I am attending", and similarly for the the speaker's notion of the hearer, **HforS**:

$\mathbf{SforH} \stackrel{\text{def}}{=} \iota a_s\,\textsc{Attending}(\mathbf{self},a_s)$

$\mathbf{HforS} \stackrel{\text{def}}{=} \iota a_h\,\textsc{Attending}(a_h,\mathbf{self})$

(we assume the standard definitions for the definite description forming operator ι). What kind of mutual knowledge results from this utterance? Well, from John's point of view (i.e., assuming that **self** is John), one has the following knowledge ordered by nesting depth n:

n

1 $\mathbf{Know}(\textsc{loves}(\mathbf{self},\mathbf{HforS}))$

2 $\mathbf{Know}(\mathbf{Know}(\mathbf{HforS},\mathbf{By}(\mathbf{SforH},$ $\textsc{loves}(\mathbf{self},\mathbf{HforS}))))$

3 $\mathbf{Know}(\mathbf{Know}(\mathbf{HforS},\mathbf{Know}(\mathbf{SforH},$ $\textsc{loves}(\mathbf{self},\mathbf{HforS}))))$

...

Note how in the $n=2$ case, the **By** operator is used to get the proposition evaluated from the speaker's perspective within the hearer's knowledge.

We define mutual indexical knowledge and associated notions as follows:

$\mathbf{MKnow}(\theta^a_{ofs},\theta^a_{sfo},\varphi) \stackrel{\text{def}}{=} \mathbf{KMK}(\theta^a_{ofs},\theta^a_{sfo},\varphi) \wedge \mathbf{By}(\theta^a_{ofs},\mathbf{KMK}(\theta^a_{sfo},\theta^a_{ofs},\mathbf{By}(\theta^a_{sfo},\varphi)))$

$\mathbf{KMK}(\theta^a_{ofs},\theta^a_{sfo},\varphi) \stackrel{\text{def}}{=} \forall n\mathbf{AKnow}(n,\theta^a_{ofs},\theta^a_{sfo},\varphi)$

$\mathbf{AKnow}(n, \theta^a_{ofs}, \theta^a_{sfo}, \varphi) \stackrel{\text{def}}{=}$
$\mathbf{Know}(\varphi)$ if $n = 1$
$\mathbf{Know}(\mathbf{Know}(\theta^a_{ofs}, \mathbf{By}(\theta^a_{sfo}, \varphi)))$ if $n = 2$
$\mathbf{Know}(\mathbf{Know}(\theta^a_{ofs}, \mathbf{By}(\theta^a_{sfo},$
$\mathbf{AKnow}(n-2, \theta^a_{ofs}, \theta^a_{sfo}, \varphi))))$ if $n > 2$

For the above example, it is easy to check that given these definitions, the formula at nesting depth n is $\mathbf{AKnow}(n, \mathbf{HforS}, \mathbf{SforH}, \textsc{loves}(\mathbf{self}, \mathbf{HforS}))$. As well, the conjunction of all nesting depths, that is, **self**'s one-sided mutual knowledge, can be expressed as $\mathbf{KMK}(\mathbf{HforS}, \mathbf{SforH}, \textsc{loves}(\mathbf{self}, \mathbf{HforS}))$. Finally, one can attribute to **self** and the hearer mutual knowledge of **self**'s love for the hearer — the conjunction of their respective one-sided mutual knowledge — by $\mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH}, \textsc{loves}(\mathbf{self}, \mathbf{HforS}))$. In general, the formula $\mathbf{MKnow}(\theta^a_{ofs}, \theta^a_{sfo}, \varphi)$ means that agents **self** and θ^a_{ofs} mutually know φ, where φ is expressed from **self**'s point of view, and where θ^a_{ofs} is the description under which **self** knows of the other agent and θ^a_{sfo} is the description under which the other agent knows of **self**.[3] The notion of mutual indexical belief could be defined in a similar manner.

Semantics of the Communication Language

A theory of linguistic communication as rational interaction of the kind proposed by Cohen and Levesque [4] must include a specification of the semantics of the communication language under consideration. If this language includes indexicals or other context sensitive elements, the semantic specification must deal with them. The most obvious way of doing this would be to directly encode a semantic account along the lines of Kaplan's [8, 7] in the logic. This would involve introducing terms that denote possible worlds and contexts into the logic and defining a truth predicate over sentences of the communication language taking such terms as arguments.

Cohen and Levesque's semantic specification in [3] is set up to handle indexicals partly along these lines: their semantic predicates TRUE, REFERS, and FULFILL-CONDS, take the utterance event as argument. So, for example, their semantics might yield

$\models$(TRUE `'I am hungry'` `e`)
$\equiv$ (SPEAKER `e` x) $\wedge$ (HUNGRY x)

where `e` is the utterance event.

How `now` could be handled in this approach is not clear as an event type cannot determine the time, but maybe this could be fixed. Another problem is that one might have to duplicate some predicates, one version taking an event as argument and the other not (e.g., (PYRAMID x `e`) and (PYRAMID x) to handle the demonstrative "this pyramid"). But a more serious problem is that the way utterance events are individuated is inappropriate if one is to account for what agents need to know to be *able* to achieve their goals by making speech acts (as in Appelt's work [1]). In the approach sketched above, utterance events determine not only the sentence used, but also who the speaker and addressee are, where the utterance is made, perhaps what gestures are made, etc. But speakers cannot be guaranteed to know all the contextual facts involved in an utterance. A speaker might not know who he's addressing, what he's pointing at, where he is, what time it is, who he is, etc. An account of ability to achieve goals by making speech acts must refer to utterance acts narrowly individuated, that is, just in terms of the sentence uttered. Now perhaps one could have both kinds of events around and relate them, but this seems unnecessarily complex.

Let us describe an alternative semantic account that does not appeal to utterance events widely individuated, and is thus more compatible with a theory of ability. Perhaps the main insight of Kaplan's work on the semantics of indexicals is that one must distinguish between the utterance context and possible worlds or indices where an expression might be evaluated. Thus, while any utterance of "I am here now" must be true, an utterance of "necessarily I am here now" is (most likely) false. As well, "I will be here 5 minutes from now" need not be true and a sincere utterance of it is quite informative. These results are due to the fact that natural language indexicals refer to whatever entity they denote in the context of utterance; this applies no matter which world or index the expression is evaluated in.

The idea behind our proposal is to interpret or translate the communication language expression into an expression in the logic while getting any indexical evaluated outside any modal context; indexicals get quantified out. So for example, for "I will be here 5 min. from now", we get

$$\models \textsc{True}(\mathtt{At}(\mathtt{now}+5, \mathtt{pos}(\mathtt{I}) = \mathtt{here})) \equiv \forall a \forall t \forall p (a = \mathbf{self} \wedge t = \mathbf{now} \wedge p = \textsc{pos}(\mathbf{self}) \supset \mathbf{At}(t+5, \textsc{pos}(a) = p))$$

where $\textsc{True}(\theta^F)$ is intended to mean that communication language sentence θ^F would be true if uttered by **self** at time **now**. Note that the formula on the right hand side is not valid, while the direct translation of $\mathtt{At}(\mathtt{now}+5, \mathtt{pos}(\mathtt{I}) = \mathtt{here})$ into the logic, $\mathbf{At}(\mathbf{now}+5, \textsc{pos}(\mathbf{self}) = \textsc{pos}(\mathbf{self}))$, is. Similarly, for "I am here now", we get:

$$\models \textsc{True}(\mathtt{pos}(\mathtt{I}) = \mathtt{here}) \equiv \forall a \forall p (a = \mathbf{self} \wedge p = \textsc{pos}(\mathbf{self}) \supset \textsc{pos}(a) = p)$$

whose right hand side is valid. But for "necessarily I am here now", we get

$$\models \textsc{True}(\mathtt{MNec}(\mathtt{pos}(\mathtt{I}) = \mathtt{here})) \equiv \forall a \forall p (a = \mathbf{self} \wedge p = \textsc{pos}(a_s) \supset \mathbf{MNec}(\textsc{pos}(a) = p))$$

[3] Note that it makes little sense to say that the agents have mutual knowledge unless **self** knows that he is θ^a_{sfo} from θ^a_{ofs}'s perspective and similarly for the other agent; perhaps this should be made part of the definition.

whose right hand side is not valid.[4]

So the idea is straightforward. What is a bit tricky is getting a compositional specification of TRUE. Let's look more closely at the "I will be here 5 min. from now" example. Notice that the expression in the logic associated with this communication language sentence consists of two parts: a set of constraints on the values of variables associated with the indexicals $a = \mathbf{self} \wedge t = \mathbf{now} \wedge p = \textsc{pos}(\mathbf{self})$ and a sentence specifying the content of the utterance in terms of these variables $\mathbf{At}(t+5, \textsc{pos}(a) = p)$. The constraints associated with the indexicals are moved out of the modal context in the content. Our compositional interpretation of communication language expressions generally follows this pattern: it concurrently builds the content and set of constraints associated with the expression. We write $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$ to specify that the content φ with constraints φ' is associated with communication language sentence θ^F. Based on $[\![\theta^F]\!]$, the semantics is specified by the following schema:

Assumption 1
$\models \textsc{True}(\theta^F) \equiv \forall V(\varphi' \supset \varphi))$,
where $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$ and
V is the set of all free variables in φ' that belong to V_i

V_i is a set of variables which may be substituted for indexicals in moving them out of modal contexts; it includes a denumerably infinite number of such variables for each sort i, a, t, and d. Also, let f be a one-one function that maps each constant that stands for a communication language variable into a variable of the appropriate sort (its interpretation) and such that no constant is mapped into an element of V_i.

The following rules specify the content and constraints associated with communication language terms:

$[\![\mathtt{I}]\!] = \langle v^a, v^a = \mathbf{self} \rangle$, $v^a \in V_i$

$[\![\mathtt{you}]\!] = \langle v^a, v^a = \mathbf{HforS} \rangle$, $v^a \in V_i$

$[\![\mathtt{now}]\!] = \langle v^t, v^t = \mathbf{now} \rangle$, $v^t \in V_i$

$[\![\mathtt{here}]\!] = \langle v^i, v^i = \textsc{pos}(\mathbf{self}) \rangle$, $v^i \in V_i$

$[\![\mathtt{V}]\!] = \langle f(\mathtt{V}), \mathbf{True} \rangle$

$[\![\mathtt{dthat}(\theta)]\!] = \langle v, \varphi \wedge v = \theta' \rangle$ iff $[\![\theta]\!] = \langle \theta', \varphi \rangle$, provided that $v \in V_i$ is not free in φ

$[\![\mathtt{f}(\theta_1, \ldots, \theta_n)]\!] = \langle \mathrm{F}(\theta'_1, \ldots, \theta'_n), \varphi_1 \wedge \ldots \wedge \varphi_n \rangle$ iff $[\![\theta_1]\!] = \langle \theta'_1, \varphi_1 \rangle, \ldots, [\![\theta_n]\!] = \langle \theta'_n, \varphi_n \rangle$ provided that $\varphi_1, \ldots, \varphi_n$ have no free variables belonging to V_i in common

For formulas, we have the following rules:

$[\![\mathtt{R}(\theta^I_1, \ldots, \theta^I_n)]\!] = \langle \mathrm{R}(\theta_1, \ldots, \theta_n), \varphi_1 \wedge \ldots \wedge \varphi_n \rangle$ iff $[\![\theta^I_1]\!] = \langle \theta_1, \varphi_1 \rangle, \ldots, [\![\theta^I_n]\!] = \langle \theta_n, \varphi_n \rangle$, provided that $\varphi_1, \ldots, \varphi_n$ have no free variables belonging to V_i in common

$[\![\theta_1 = \theta_2]\!] = \langle \theta'_1 = \theta'_2, \varphi_1 \wedge \varphi_2 \rangle$ iff $[\![\theta_1]\!] = \langle \theta'_1, \varphi_1 \rangle$ and $[\![\theta_2]\!] = \langle \theta'_2, \varphi_2 \rangle$, provided that φ_1 and φ_2 have no free variables belonging to V_i in common

$[\![\mathtt{Does}(\theta^D, \theta^T)]\!] = \langle \mathbf{Does}(\theta^d, \theta^t), \varphi_d \wedge \varphi_t \rangle$ iff $[\![\theta^D]\!] = \langle \theta^d, \varphi_d \rangle$ and $[\![\theta^T]\!] = \langle \theta^t, \varphi_t \rangle$, provided that φ_d and φ_t have no free variables belonging to V_i in common

$[\![\neg\theta^F]\!] = \langle \neg\varphi, \varphi' \rangle$ iff $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$

$[\![\theta^F_1 \wedge \theta^F_2]\!] = \langle \varphi_1 \wedge \varphi_2, \varphi'_1 \wedge \varphi'_2 \rangle$ iff $[\![\theta^F_1]\!] = \langle \varphi_1, \varphi'_1 \rangle$ and $[\![\theta^F_2]\!] = \langle \varphi_2, \varphi'_2 \rangle$, provided that φ'_1 and φ'_2 have no free variables belonging to V_i in common

$[\![\forall \mathtt{V}\, \theta^F]\!] = \langle \forall f(\mathtt{V})\, \varphi, \varphi' \rangle$ iff $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$

$[\![\mathtt{At}(\theta^T, \theta^F)]\!] = \langle \mathbf{At}(\theta, \varphi), \varphi_\theta \wedge \varphi' \rangle$ iff $[\![\theta^T]\!] = \langle \theta, \varphi_\theta \rangle$ and $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$, provided that φ_θ and φ' have no free variables belonging to V_i in common

$[\![\Box\theta^F]\!] = \langle \Box\varphi, \varphi' \rangle$ iff $[\![\theta^F]\!] = \langle \varphi, \varphi' \rangle$

The rules for $\theta^T_1 < \theta^T_2$ and $\mathtt{By}(\theta^A, \theta^F)$ are similar to that for $\theta_1 = \theta_2$ and $\mathtt{At}(\theta^T, \theta^F)$ respectively. The interpretation assigned to the examples can be checked to follow from these rules. We would like to extend this to handle attitude reports. This raises difficult problems though. In this case the strategy of moving indexicals out of modal contexts only yields *de re* readings. One would also want an account of why some non-*de re* readings and not others are suggested by such utterances depending on the context.

Communicative Acts and Ability

In this section, we develop a simple account of the pragmatics of declaratives in the style of Appelt [1]. The account uses the communication language semantics of the previous section, and thus can deal with indexicality in the communication language as well as in thought and action. The account also allows one to model (and reason about) the ability of agents to achieve goals by making utterances. Note that the account is quite simplistic in that it ignores the possibility of insincerity, irony, indirect speech acts, etc.

We formalize declaratives as follows. Our first assumption says that it is physically possible for **self** to make a declarative utterance of sentence θ^F if and only if he is addressing a unique hearer who is attending only to him, he knows that the sentence is true, and he wants the hearer to know that the sentence is true from the speaker's perspective:

Assumption 2 (Preconditions of UTTRDCL**)**
$\models \mathbf{PhyPoss}(\textsc{uttrDcl}(\theta^F)) \equiv$
$\forall a(a = \mathbf{self} \supset \mathbf{By}(\mathbf{HforS}, \mathbf{SforH} = a)) \wedge$
$\mathbf{Know}(\textsc{True}(\theta^F)) \wedge$
$\mathbf{Goal}(\mathbf{Know}(\mathbf{HforS}, \mathbf{By}(\mathbf{SforH}, \textsc{True}(\theta^F))))$

Note that it would be better to talk about belief rather than knowledge in this formalization. However, ability to achieve goals by doing complex actions requires that

[4] **MNec** stands roughly for metaphysical necessity. $\mathbf{MNec}(\varphi) \stackrel{\text{def}}{=} \forall v^t \mathbf{At}(v^t, \Box\varphi)$, where v^t does not occur free in φ, that is, **MNec** stands for "always historically necessary"; this approximation of metaphysical necessity is good enough for our purposes.

the agent remember relevant facts about the initial situation and the results of previous actions; in the case of knowledge, this can be taken care of by assuming that knowledge is persistent; but the analogous assumption for belief would be much less reasonable; an account of belief revision would be needed. It would also be better to talk about intention rather than **Goal**; we could borrow Cohen and Levesque's formalization of intention [2] to provide for that. We will try to alleviate these deficiencies in future versions of the account.

Our second assumption says that if **self** and the hearer mutually know that **self** is addressing a unique agent who is attending only to him, then after **self** utters declarative sentence θ^F, it must be the case that **self** and the hearer have the mutual indexical knowledge that **self** has just uttered the sentence:[5]

Assumption 3 (Effects of $\textsc{uttrDcl}$**)**

$$\begin{aligned}&\models \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH},\\ &\qquad \forall a(a = \mathbf{self} \supset \mathbf{By}(\mathbf{HforS}, \mathbf{SforH} = a))) \supset\\ &\quad \mathbf{AfterNec}(\textsc{uttrDcl}(\theta^F),\\ &\qquad \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH}, \mathbf{Done}(\textsc{uttrDcl}(\theta^F))))\end{aligned}$$

We also assume that agents always know how to make declarative utterances:

Assumption 4 ($\textsc{uttrDcl}$ **is known)**

$$\models \exists d \mathbf{Know}(d = \textsc{uttrDcl}(\theta^F))$$

From these assumptions, one can then show that if **self** and the hearer mutually know that **self** is addressing a unique agent who is attending only to him, and **self** knows that sentence θ^F is true and wants the hearer to know that θ^F is true from the speaker's perspective, then by making a declarative utterance of θ^F, **self** is able to achieve the goal that he and the hearer have mutual indexical knowledge that when the sentence was uttered, the speaker knew that the sentence was true and wanted the hearer to know that it was true from the speaker's perspective:

Proposition 1

$$\begin{aligned}&\models \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH},\\ &\qquad \forall a(a = \mathbf{self} \supset \mathbf{By}(\mathbf{HforS}, \mathbf{SforH} = a))) \wedge\\ &\quad \mathbf{Know}(\textsc{True}(\theta^F)) \wedge\\ &\quad \mathbf{Goal}(\mathbf{Know}(\mathbf{HforS}, \mathbf{By}(\mathbf{SforH}, \textsc{True}(\theta^F)))) \supset\\ &\quad \mathbf{Can}(\textsc{uttrDcl}(\theta^F),\\ &\qquad \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH},\\ &\qquad\quad \mathbf{DoneWhen}(\textsc{uttrDcl}(\theta^F),\\ &\qquad\qquad \mathbf{Know}(\textsc{True}(\theta^F)) \wedge\\ &\qquad\qquad \mathbf{Goal}(\mathbf{Know}(\mathbf{HforS}, \mathbf{By}(\mathbf{SforH}, \textsc{True}(\theta^F)))))))\end{aligned}$$

If we instantiate this for the sentence "I will be here 5 min. from now", we get that if it is mutually known among **self** and the hearer that **self** is addressing a unique agent who is attending only to him, and **self** knows that he will be at his then current position 5 min. later and wants the hearer to know it too, then by uttering the sentence, **self** is able to achieve the goal that he and the hearer mutually know that when the sentence was uttered, the speaker knew that he would be at his then current position 5 min. later and wanted the hearer to know that too:

Proposition 2

$$\begin{aligned}&\models \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH},\\ &\qquad \forall a(a = \mathbf{self} \supset \mathbf{By}(\mathbf{HforS}, \mathbf{SforH} = a))) \wedge\\ &\quad \mathbf{Know}(\varphi) \wedge \mathbf{Goal}(\mathbf{Know}(\mathbf{HforS}, \mathbf{By}(\mathbf{SforH}, \varphi))) \supset\\ &\quad \mathbf{Can}(\textsc{uttrDcl}(\mathtt{At}(\mathtt{now} + 5, \mathtt{pos}(\mathtt{I}) = \mathtt{here})),\\ &\qquad \mathbf{MKnow}(\mathbf{HforS}, \mathbf{SforH},\\ &\qquad\quad \mathbf{DoneWhen}(\textsc{uttrDcl}(\mathtt{At}(\mathtt{now} + 5, \mathtt{pos}(\mathtt{I}) = \mathtt{here})),\\ &\qquad\qquad \mathbf{Know}(\varphi) \wedge\\ &\qquad\qquad \mathbf{Goal}(\mathbf{Know}(\mathbf{HforS}, \mathbf{By}(\mathbf{SforH}, \varphi))))))\end{aligned}$$

where $\varphi \stackrel{\text{def}}{=}$

$$\begin{aligned}&\forall a \forall t \forall p(a = \mathbf{self} \wedge t = \mathbf{now} \wedge p = \textsc{pos}(\mathbf{self}) \supset\\ &\qquad \mathbf{At}(t + 5, \textsc{pos}(a) = p))\end{aligned}$$

Note that the knowledge involved in this example is indexical; in particular, neither agent need know where he is or what time it is. It should be possible to extend this example and show that under the right conditions, the agent can utter the sentence, go take care of some business, and then return and meet with the hearer.

Conclusion

In this paper, we have presented a framework where the indexicality of language can be modeled and related to the indexicality of mental states and action. Our semantics for the indexical communication language does not require terms denoting contexts to be introduced. And our formalization of the pragmatics of declarative utterances allows one to model the ability of agents to achieve goals by making such utterances. This research can be seen as providing a more accurate specification of the reasoning that a communicating agent (natural language interface, DAI system, etc.) must do to be coherent and helpful.

In future work we plan to develop the framework to the point where it can handle complex linguistic and physical interaction between agents and their environment. Our account of ability needs to be extended to handle plans involving multiple agents in a general way. A formalization of the relationship between intention, ability, and the eventual achievement of goals is also needed. We need to verify that the semantics can handle a wide range of indexical expressions. Some aspects of the design of the logic, the behavior of **self** and **now** in particular, should perhaps be revised to simplify the mapping of the communication language into the logic. The current design favors simplicity and conciseness at the expense of a simple relationship to natural language. We also want to develop a more realistic version of our pragmatics, handle other types of speech acts, performatives, etc. Finally, we feel that other issues at the intersection of semantics and prag-

[5] $\mathbf{Done}(\delta) \stackrel{\text{def}}{=} \mathbf{DoneWhen}(\delta, \mathbf{True})$. $\mathbf{DoneWhen}(\delta, \varphi)$ means that **self** has just done δ and that φ was true when he started; $\mathbf{DoneWhen}(\delta, \varphi) \stackrel{\text{def}}{=} \exists v_s^t \exists v_e^t (v_e^t = \mathbf{now} \wedge \mathbf{At}(v_s^t, \mathbf{Does}(\delta, v_e^t) \wedge \varphi))$, provided that v_s^t and v_e^t are distinct and do not occur anywhere in φ or δ.

matics, such as reference, would gain to be examined within a unified framework such as ours.

Acknowledgement

I want to thank Dr. Yasuhiro Katagiri for sponsoring my stay at NTT Basic Research Laboratories, where I started working on the ideas developed in this paper. Discussions with him and other members of the Takeuchi group, as well as with Hector Levesque and Phil Cohen have been helpful.

References

[1] Douglas E. Appelt. *Planning English Sentences.* Cambridge University Press, Cambridge, UK, 1985.

[2] Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

[3] Philip R. Cohen and Hector J. Levesque. Performatives in a rationally based speech act theory. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, pages 79–88, Pittsburg, PA, June 1990.

[4] Philip R. Cohen and Hector J. Levesque. Rational interaction as the basis for communication. In Philip R. Cohen, J. Morgan, and Martha E. Pollack, editors, *Intentions in Communication.* MIT Press, Cambridge, MA, 1990.

[5] Robert Goldblatt. *Logics of Time and Computation.* CSLI Lecture Notes No. 7. Center for the Study of Language and Information, Stanford University, Stanford, CA, 1987.

[6] Jaakko Hintikka. Semantics for the propositional attitudes. In J. W. Davis, D. J. Hockney, and K. W. Wilson, editors, *Philosophical Logic.* D. Reidel Publishing, Dordrecht, Holland, 1969.

[7] David Kaplan. On the logic of demonstratives. In Peter A. French, Theodore D. Jr. Uehling, and Howard K. Wettstein, editors, *Contemporary Perspectives in the Philosophy of Language*, pages 401–412. University of Minnesota Press, Minneapolis, 1979.

[8] David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan.* Oxford University Press, Oxford, 1989.

[9] Saul A. Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16:83–94, 1963.

[10] Yves Lespérance. *A Formal Theory of Indexical Knowledge and Action.* PhD thesis, Department of Computer Science, University of Toronto, January 1991. Also published as technical report CSRI-248.

[11] Yves Lespérance and Hector J. Levesque. On the logic of indexical knowledge and action. In preparation.

[12] Yves Lespérance and Hector J. Levesque. Indexical knowledge in robot plans. In *Proceedings of the Eight National Conference on Artificial Intelligence*, pages 868–874, Boston, August 1990. American Association for Artificial Intelligence, AAAI Press/MIT Press.

[13] Yves Lespérance and Hector J. Levesque. Indexical knowledge and robot action — a logical account. Submitted to *Artificial Intelligence*, 1992.

[14] Robert C. Moore. Reasoning about knowledge and action. Technical Report 191, AI Center, SRI International, Menlo Park, CA, October 1980.