What do you mean by "what do you mean"?: A formal representation and its dynamic semantics of meta-expressions in conversations

Norihiro Ogata

Faculty of Language and Culture, Osaka University
1-8 Machikaneyama, Toyonaka, Osaka, Japan
email:ogata@lang.osaka-u.ac.jp

Abstract. This paper investigates dynamic semantics of conversations from the point of view of semantical closedness, presuppositions and shared belief/common knowledge updates, by analyzing the meta-expression "what do you mean (by X)?" into three major usages: *semantic repair initiation*, *intentional repair initiation*, and *inferential repair initiation*, since these three usages are deeply related with three types of semantical closedness: *closedness of denotations*, *closedness of intention* and *closedness of inference* of conversations. As a result, the proposed dynamic semantics of conversations is semantically closed in terms of shared beliefs of the conversants.

1 Introduction

A conversation can be regarded as a cooperative social action that forms and updates the common knowledge of the society, community, or group. Therefore, to answer the question of how conversations can form or update the common knowledge of the community is to propose dynamic semantics of conversations. As a step towards proposing dynamic semantics of conversations, we will concentrate on the meaning of the phrase "what do you mean (by X)?". This phrase can be interpreted in at least three ways that are deeply relevant to forming or updating the commn knowledge of the community. This means that analyses of this phrase can give a hint of modelling of formations and updates of the common knowledge of a community and conversants' commitment to the processes.

In section 2, an analysis of three interpretations of the phrase "what do you mean (by X)?" will be shown, and a formal language modelling conversations will be defined. In section 3, dynamic semantics of conversation will be defined.

2 *What do you mean* means three types of triggers of updates of common knowledge

The phrase "what do you mean (by X)?", which functions an initiation of repairs in the sense of conversational analysis ([9],etc.), is a typical phrase used in conversation. A consideration of its semantics is useful for proposing semantics of conversations, since the phrase has at least three uses:

(I) asking what the phrase X denotes (*semantic repair initiation*),
(II) asking what the previous speaker intends by saying the phrase X (*intentional repair initiation*), and
(III) asking what grounds of the previous speaker's statement (*inferential repair initiation*),

and all three meanings are deeply relevant to the formation or update of the common knowledge of a community, since (I) implies that in the community the denotation of X is not defined as common knowledge but as the previous speaker's private knowledge, (II) implies that in the community the implication relation between X and its intended content is not shared, and (III) implies that in the community the deduction of X from some bases is not shared.

We can find examples of all three interpretations of the phrase in documents distributed over the WWW, as follows:

(1) a. A: ... Because the country is so different today from then. It's a different world entirely. Then you didn't have wages or crisis...
B: And *what do you mean by then*?
A: Back in the 30s. Let's take 60 years.
B: Okay.

b. A: *What do you mean by* saying that our sorrow should be interior?
B: When I say that our sorrow should be interior,
I mean that it should come from the heart,
and not merely from the lips.

c. A: ... Why do you want the dog to stop?
it's not perverted - it's dog language. ...
B: *What do you mean* why do I want it to stop???
It's freaking embarassing if guests come over and my dog humps them.
I do not want it.

(1a-c) can be represented in formal schemata, as follows:

(2) a. $p[x]; ?y.denotes(x, y)$
b. $p; ?q.p \Rightarrow q$
c. $p; ?y.y \vdash p$

where $p[x]$ is an atomic move which transmits the proposition p with an argument x, ";" means a turn exchange (i.e., conversational dynamic conjunction in the sense of dynamic semantics [10, 7]), $?x.p$ means "in the previous expression p, what does x means?", $denotes(x, y)$ means x denotes y, $\Rightarrow$ means an implication, $x_1, \ldots, x_n \vdash p$ means that p is deducible from the context $x_1, \ldots, x_n$.

Since, by the deduction theorem, $x_1, \ldots, x_n \vdash p$ is equivalent to $\vdash x_1 \wedge \ldots \wedge x_n \Rightarrow p$, (2c) is reformulated as follows:

(3) $p; ?y.y \Rightarrow p$

These formulas (i.e., (2a-b) and (3)) are defined by a BNF-grammar as in the following definition.

Definition 1.

$$v \in Var_{INDIV} \qquad p \in Var_{PROP}$$
$$\tau \in TERM \qquad \phi \in \Phi$$
$$\varphi \in FORM_{\Phi} \qquad m \in MOVE_{\Phi}$$
$$\phi ::= \pi(\tau_1, \ldots, \tau_n) | denotes(\tau_1, \tau_2) | \tau_1 = \tau_2$$
$$\varphi ::= \phi | \varphi_1 \Rightarrow \varphi_2 | \varphi_1 \wedge \varphi_2 | \bot$$
$$m ::= \varphi | ?x.\varphi | agree | m_1; m_2$$

where Var_{INDIV} is a set of individual variables, Var_{PROP} a set of propositional variables, $TERM$ a set of terms, Φ a set of atomic formulas, $FORM_{\Phi}$ a set of formulas, $MOVE_{\Phi}$ a set of sequences of moves, $\pi \in PRED$ an n-place predicate, $denotes(\tau_1, \tau_2)$ means that τ_1 denotes τ_2, x is a propositional variable or an individual variable, and $agree$ means an assent move.

This language contains a semantic predicate *denotes* and an implication $\Rightarrow$ which means both of an implication and a deductive relation. That is, this language is semantically, intentionally and inferentially closed in the sense that the denotation of terms, intention, which is treated as an implication, and inference, which is also treated as an implication, can be described within the language.

Intuitively, the semantics of each move is defined as a kind of conditional, since its semantics cannot be defined independently of its contenxt as the following observation:

(4) a. $p[x]; agree$ iff $(denotes(x, y) \wedge p[x] \wedge p[y])$ is shared in the community ... (2a)
b. $p; agree$ iff $((p \Rightarrow y) \wedge p \wedge y)$ is shared in the community ... (2b)
c. $p; agree$ iff $((\Gamma \Rightarrow p) \wedge \Gamma \wedge p)$ is shared in the community ... (3)

That is, assertive move p can change the community's shared knowledge if it is accepted in the community, and for the acceptance in case (2a) $denotes(x, y)$ must be shared as its common presupposition, in case (2b) $p \Rightarrow y$ for some "y" in the context, and in case (3) $\Gamma \Rightarrow p$ for some context Γ, respectively.

Conversely, when we say "what do you mean (by X)?", we consider such presuppositions as possibly undefined, and this means the phrase pragmatically urges re-sharing of such presuppositions, or testing of sharing of the presuppositions.

To sum up, the move p means a conditional "if the presupposition of p for its understanding is defined or shared in the community, update the common knowledge of the community by adding proposition what p means to the common knowledge", and "what do you mean (by X)?" means a declaration of undefinedness of such presuppositions.

3 A Dynamic Semantics of Conversations

More formally, exploiting the idea of update semantics [11], Discourse Representation Theory-based dynamic semantics of dialogue [8], and dynamic epistemic semantics ([4,3,2]), dynamic semantics of each sequence of moves m is defined in the basis of conditional updates discussed in section 2.

Definition 2. *Given a Kripke model of an epistemic logic* $(W, (R_a)_{a \in Agent}, (D_w)_{w \in W}, I)$, *where* W *is a set of states,* $Agent$ *a set of agents,* $R_a \subseteq W \times W$, D_w *is the set of individuals in state* w, *and* $I : W \times TERM \to D_w \cup W \times PRED \to pow(D_w^n)$, *a dynamic interpretation of* $MOVE_\Phi$ $[\![,]\!] : CONTEXT \times W \to CONTEXT \times W$ *is defined by induction of the complexity of* $m \in MOVE_\Phi$, *as follows:*

- $[\![\varphi]\!](c, X) = (c', P)$, *where* $c' = c[\varphi/c(prev), P/c(prevw)]$, $P = \{w \in Y | Presup(\varphi, w, c) = ! \text{ implies } w \models shared(c(cm), \varphi)\}$ *and if* $top(c(world)) \neq \Lambda$, *then* $Y = top(c(world))$ *and* $c'(world) = pop(c(world))$, *otherwise* $Y = X$,
- $[\![?x.\varphi]\!](c, X) = (c', P)$, *where* $P = \{w \in X | Presup(\varphi, w, c) = ?\})$ *and* $c' = c[\varphi/c(prev), push(c(prev), c(foc))/c(foc), P/c(prevw), push(c(world), c(prevw))/c(world)]$,
- $[\![agree]\!](c, X) = (c', \{w \in X | Presup(top(c(foc)), w, c) = !\})$, *where* $c' = c[pop(c(foc))/foc(c)]$,
- $[\![m_1; m_2]\!](c, X) = [\![m_2]\!](ex(c), [\![m1]\!](c, X))$.

where c *is a contextual index and* $ex(X)$ *are operations of contextual indices,* $push(X)$ *and* $pop(X)$ *are operations of the focus stack of a context, defined in definition 3,* $Presup(X, Y, Z)$ *are also defined in definition 4, and* $w \models X$ *(*X *is true at* w*), meta-languages* $shared(cm, Y)$ *(*Y *is shared among the members of the community* cm*) is defined in definition 5.*

Definition 3. $c \in CONTEXT = \{cm, spk, hrr, foc\} \hookrightarrow Agent \cup pow(Agent) \cup Stack(FORM_\Phi)$ *is a contextual index satisfying the following conditions:*

- $c(cm) \subseteq Agent$, *i.e., the community in the context* c,
- $c(spk) \in Agent$ *and* $c(hrr)$ *are the speaker and the hearer in the context* c, *where* $c(spk) \neq c(hrr)$, *respectively,*
- $c(prevw) \in pow(W)$, $c(world) \in Stack(pow(W))$,
- $c(prev) \in FORM_\Phi$, $c(foc) \in Stack(FORM_\Phi)$,
- $ex : CONTEXT \to CONTEXT$ *such that* $spk(ex(c)) = c(hrr)$, $hrr(ex(c)) = c(spk)$, *and* $ex(c(foc)) = c(foc)$,

where $Stack(X)$ *is a class satisfying the following conditions:*
$\Lambda \in Stack(X)$,
if $s \in Stack(X)$ *and* $x \in X$ *then* $push(s, x) \in Stack(x)$,
if $s \in Stack(X)$, *then* $top(push(s, x)) = x$,
if $s \in Stack(X)$, *then* $pop(push(s, x)) = s$.

We can handle nested conversations such as (5) by the focus stack in the above definition.

(5) A: Hallowed be thy name..... B: Hold it. *What do you mean by that*? A: *By what*? B: By "hallowed be thy name"? A: It means.....

Definition 4. $Presup : (FORM_\Phi \cup TERM_\Phi) \times W \times CONTEXT \to \{!, ?\}$ *('!' means 'defined' and '?' 'undefined') is a check function of presuppositions*[1] *of expressions at a state and a context which is shared among the members of the community, and defined by induction on the conplexity of* φ *as follows:*

- $Presup(\tau, w, c) = !$ *iff* $w \models shared(c(cm), \tau = \tau')$ *and* $Presup(\tau', w, c) = !$,

[1] This meaning of the term 'presuppositions' has a sense wider than a usual sense of the term which can be found in [5, 1, 6].

- $Presup(denotes(\tau_1, \tau_2), w, c) =!$ *iff* $w \models shared(c(cm), \tau_1 = \tau_2)$ *and* $Presup(\tau_2, w, c) =!$,
- $Presup(\pi(\tau_0, \ldots, \tau_{n-1}), w, c) =!$ *iff for all* $i < n$, $Presup(\tau_i, w, c) =!$ *and* $Ground(\pi(\tau_0, \ldots, \tau_{n-1}), w, c) =!$,
- $Presup(\varphi_1 \Rightarrow \varphi_2, w, c) =!$ *iff* $Ground(\varphi_1 \Rightarrow \varphi_2, w, c) =!$, $Presup(\varphi_1, w, c) =!$ *and* $Presup(\varphi_2, w, c) =!$,
- $Presup(\varphi_1 \wedge \varphi_2, w, c) =!$ *iff* $Presup(\varphi_1, w, c) =!$ *and* $Presup(\varphi_2, w, c) =!$,
- $Ground(\varphi, w, c) =!$ *iff* $w \models shared(c(cm), \varphi)$ *or* $((w \models shared(c(cm), p \Rightarrow \varphi)$ *or* $w \models shared(c(cm), \varphi \Rightarrow p))$ *and* $Presup(p, w, c) =!)$.

The definition of $Presup(denotes(\tau_1, \tau_2), w, c)$ in the above definition means that τ_1 is defined as the other term which is already shared in the community, that is, *denotes* is defined not as a purely semantic predicate, but as a kind of *renaming* predicate.

Definition 5. *Given a Kripke model of an epistemic logic* $(W, (R_a)_{a \in Agent}, (D_w)_{w \in W}, I)$, $w \models \varphi$ *means that* p *is true at* $w \in W$, *satifying the following conditions:*

- $w \not\models \bot$
- $w \models \pi(\tau_1, \ldots, \tau_n)$ *iff* $(I_w(\tau_1), \ldots, I_w(\tau_n)) \in I_w(\pi)$
- $w \models denotes(\tau_1, \tau_2)$ *iff* $I_w(\tau_1) = I_w(\tau_2)$
- $w \models \tau_1 = \tau_2$ *iff* $I_w(\tau_1) = I_w(\tau_2)$
- $w \models \varphi_1 \wedge \varphi_2$ *iff* $w \models \varphi_1$ *and* $w \models \varphi_2$
- $w \models \varphi_1 \Rightarrow \varphi_2$ *iff* $w \models \varphi_1$ *implies* $w \models \varphi_2$
- $w \models shared(cm, \varphi)$ *iff the largest fixed point of equation*

$$q = \bigcap_{a \in cm} Bel(a)([\![\varphi]\!] \cap q)$$

where $[\![\varphi]\!] = \{w \in W | w \models \varphi\}$, *and*

$$Bel(a)(X) = \{w \in W | wR_a u \text{ implies } u \in X\}.$$

The following example is the interpretation of the formalized conversation of a conversation such as (1a).

Example 1.

$$\begin{aligned}
&[\![\pi(a,b); ?x.denotes(b,x); denotes(b,d); agree]\!](c, X) \\
&= [\![agree]\!]([\![denotes(b,d)]\!](c', \{w | Presup(denotes(b,x), c, w) =?, P_1(w)\}) \\
&= (c'', \{w \in top(c'(world)) | Presup(denotes(b,d), c', w) =!, w \models shared(c'(cm), denotes(b,d))\})
\end{aligned}$$

where $c' = c[push(pop(foc(c)), \pi(a,b))/foc(c), push(\{w \in X | P_1(w)\}, c(world))/c(world), ...]$, $c'' = c'[push(pop(foc(c)), denotes(b,d))/foc(c)$, $push(\{w \in top(c'(world)) | P_2(w)\}, c(world))/c(world), ...]$, $P_1(w) = Presup(\pi(a,b), c, w) = ! \rightarrow w \models shared(c(cm), \pi(a,b))$, and $P_2(w) = Presup(denotes(b,d), c', w) =! \rightarrow w \models shared(c'(cm), denotes(b,d))$.

In this conversation, if intentional or inferential presuppositions of the first move are shared, the content of the first move is shared, and this property is reflected in example 1.

4 Conclusion

We have discussed the meaning of "what do you mean (by X)?" and its dynamic semantics based on updates of the common knowledge of the community. To clarify the meaning of "what do you mean (by X)?", a new concept of presuppositions has been defined in terms of shared beliefs among the members of the conversation. As a result, we could propose dynamic semantics and representation that can grasp the semantically-closed or meta-linguistic closed character of conversation, i.e., closedness on denotation and inference. Furthermore, we have been able to treat nested structures of conversations introduced by "what do you mean (by X)?" and other repair initiation moves by defining contexts by a kind of stack structure.

References

1. David Ian Beaver. Presupposition. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*. Elsevier Science B. V., Amsterdam, 1997.
2. Jelle Gerbrandy. Dynamic epistemic logic. In Lawrence S. Moss, Jonathan Ginzburg, and Maarten de Rijke, editors, *Logic, Language and Computation: vol. 2*, pages 67–84. CSLI Publications, Stanford, 1999.
3. Jelle Gerbrandy and Willem Groeneveld. Reasoning about information change. *Journal of Logic and Computation*, 23:267–306, 1997.
4. Willem Groeneveld. *Logical Investigations into Dynamic Semantics*. PhD thesis, University of Amsterdam, 1995. published by ILLC-Publications.
5. Irene Heim. On the projection problem for presuppositions. In D. Flickinger et al., editors, *Proceedings of the Second West Coast Conference on Formal Linguistics*, pages 114–125. Stanford, 1988.
6. Emiel Krahmer. *Presupposition and Anaphora*. CSLI Publications, Stanford, 1998.
7. R. Muskens, J. van Benthem, and A. Visser. Dynamics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 587–648. Elsevier Science B.V., Amsterdam, 1997.
8. Norihiro Ogata. Formal semantics of dialogues based on belief sharing and observational equivalence of dialogues. *Journal of Natural Language Processing*, 6(4):93–115, 1999.
9. Emanuel A. Schegloff, Gail Jefferson, and Harvey Sacks. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382, 1977.
10. Johan van Benthem. *Exploring Logical Dynamics*. CLSI Publications, Stanford, 1996.
11. Frank C. Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 8:73–80, 1996.