Information Update by Corrections

Sumiyo Nishiguchi[1,a)]

Abstract: This paper is going to discuss belief update in cases of corrections. While van Benthem (2000) discusses public announcement of what the agent had been ignorant of, which involves elimination of possible worlds, corrections entail elimination and addition of possible worlds to one's belief. One's belief might be corrected by means of elimination of believed worlds and addition of the possible worlds that are claimed to be true. Whereas assertion of negative sentences eliminates p-worlds and adds non-p-worlds on one's belief worlds, a contrary proposition q eliminates p-worlds and adds q-worlds.

1. Belief Update as World Elimination

Considering belief update as world elimination is not a new idea. van Benthem (2000) discusses public update as world elimination. When Mary is not sure if p is true or false, public announcement of p eliminates $\neg p$ worlds from Mary's belief worlds.

In van Benthem (2000), the factual information of a single agent is represented by a set of possible worlds which might still be the actual world as far as the agent knows. When the agent is uncertain regarding p or $\neg p$ as in (1), the agent's information is updated by the assertion p which eliminates $\neg p$ worlds as illustrated in (2).

(1)

p	¬ p

(2)

p

For a group of agents, public announcement of p eliminates $\neg p$ worlds from the common knowledge. For example, consider a case when two people engage in conversation. When an agent A asks the other person B a question p and B answers "yes," all $\neg p$ worlds are removed and the agents reach common knowledge of p.

2. Information Update by Corrections

While van Benthem (2000) discusses world elimination of the knowledge of an ignorant agent who is not sure either p or $\neg p$ holds, this paper considers the contribution of corrections of prior belief executed by discourse participants. I claim that corrections entail elimination and addition of possible worlds to one's belief. One's knowledge might be corrected by means of elimination of believed worlds and addition of possible worlds that are claimed to be true.

[1] Tokyo University of Science, Kuki, Saitama 346–8512, Japan
[a)] nishiguchi@rs.tus.ac.jp

Information update adds more propositions that are true to one's belief worlds. We will define an update system for a language L to be <L, Σ, [ ] > building on Veltman (1996). Σ is a set of relevant information states, a function [ ] assigns each sentence ϕ and operation $[\phi]$ on Σ. $\sigma[\phi]$ is the result of updating σ with ϕ. The information state σ is a set of possible worlds which the agent x believes to be true.

(3) Let σ be an information state, w be a possible world, x be an agent, p, q, r be propositions.

$\sigma = \lambda$w.[B(p(w) = q(w) = r(w) = 1)(x)]

(B: believe)

The possible worlds which the person believes to be true are intersection of believed propositions.

(4) $\sigma = \lambda$w.[B(the sun rises from east(w) ∧ Noda is the prime minister of Japan(w) ∧ ice is made from water(w),...)(x)]

When one keeps acquiring new propositions believed to be true, the number of the possible worlds are supposed to be reduced.

Learning falsity of one's belief and receiving corrections undergo eliminative process. For example, one might have believed that Kan was the current prime minister of Japan but now someone corrects his belief.

(5) $\sigma = \lambda$w.[B(Kan is the prime minister of Japan(w) ∧ the sun rises from the east(w) ∧ Bob is x's neighbor(w))(x)]

$\sigma[\phi] = \lambda$w.[B(the sun rises from the east(w) ∧ Bob is x's neighbor(w))(x)]

ϕ = Kan is not the prime minister of Japan

The number of propositions x believes is now one less than what x used to believe before correction. Acceptance of correction results in elimination of propositions out of common ground.

What if the correction was made by the utterance "Noda is the prime minister of Japan"? Even without a negative marker, elimination of the proposition "Kan is the prime minister of Japan"

is achieved and a new proposition "Noda is the prime minister of Japan" is added to x's belief. The next section will further discuss these two types of corrective sentences.

3. Two Types of Corrective Sentences

In (5), the corrective assertion ϕ is a negative sentence, that is, a predicate negation or nexal negation of the proposition *Kan is the prime minister of Japan* (Jespersen 1917). Correction can be done by negative sentences or by sentences which express contrary propositions. There could be two types of corrective sentences:

(1) Assertion of a negative sentence whose proposition is $\neg p$ eliminates p-worlds and adds $\neg p$-worlds on one's belief.

(2) Assertion of a contrary proposition q eliminates p-worlds and adds q-worlds.

For example, the negative counterpart of the sentence whose proposition is p, is $\neg p$. The negative sentence contains the negative *not* which contributes to predicate negation. On the other hand, the sentence form of contrary proposition q contains the antonym of *west* which is *east* (*contrariety* in Aristotle, cf.Horn 2001).

(6) p: The sun rises from the west.

$\neg p$: The sun does not rise from the west (it could rise from either of other directions).

q: The sun rises from the east.

Although contrary propositions entail contradictory propositions, contradictory propositions contain what is called *excluded middle* which do not belong to either p or q. For example, $\neg p$ may contain worlds in which the sun neither rises from the east or the west—the sun may rise from the north or the south in $\neg p$ worlds. Therefore, q worlds are the subset of $\neg p$ worlds.

(7) a. $q \models \neg p$

b. The sun rises from the east $\models$ The sun does not rise from the west

Assertion of the interpreted q-sentences eliminates p-worlds and adds q worlds, which entails adding $\neg p$ worlds.

4. Temporal Dynamic Semantics

In the present state of temporal dynamic semantics (Nishiguchi 2003, 2007), which builds on Veltman (1996) and is hinted by Condoravdi (2002), information update of common belief by corrective assertions can be formulated as follows:

(8) a. Correction by a contradictory assertion $\neg p$:

PAST B(p)(x) $\wedge$ PRES B($\neg$p)(x$\oplus$y)

b. Correction by a contrary assertion q:

PAST B(p)(x) $\wedge$ PRES B(q)(x$\oplus$y)

(PAST: past, PRES: present, x, y: discourse participants, $\alpha \oplus \beta$: individual sum of α and β à la Link 1983)

Corrective assertions presuppose the hearer's past belief on p which is replaced by the present common belief in $\neg p$ in (8a). On the other hand, contrary assertion overwrites the hearers past belief in p with the present shared belief in q as in (8b).

5. Conclusion

This paper investigated contributions of corrective assertions. Corrective assertions can be made either by contradictory propositions or contrary propositions. The former eliminates the possible worlds in which p is true out of the agent's belief worlds and adds $\neg p$ worlds. The latter deletes the belief in p and adds q-worlds which logically entails the belief in $\neg p$.

References

[Aristotle] Aristotle, Owen, O.F. and Porphyry: 1901, 2009, *The Organon; Or Logical Treatises of Aristotle: With the Introduction of Porphyry*, General Books.

[van Benthem 2000] van Benthem, J.: 2000, Information update as relativisation. ILLC, University of Amsterdam, Available from http://staff.science.uva.nl/~johan/.

[Condoravdi 2001] Condoravdi, C.: 2001, Temporal interpretation of modals. *The Construction of Meaning*, 59-87, CSLI, Stanford.

[Horn 2001] Horn, L. R.: 2001, *A Natural History of Negation*, CSLI, Stanford.

[Jespersen 1917] Jespersen, O.: 1917, *Negation in English and Other Languages*, Høst, København.

[Link 1983] Godehard, Link:. The logical analysis of plurals and mass terms: a lattice-theoretical approach, *Meaning, Use and Interpretation of Language*, 303-323, 1983, Walter de Gruyter, Berlin, New York.

[Nishiguchi 2004] Nishiguchi, S.: 2004, Temporal dynamic semantics of factual counterfactuals, *Proceedings of JSAI 18 International Workshop LENLS 2004*, 73-81, Kanazawa, Japan.

[Nishiguchi 2007] Nishiguchi, S.: 2007, Temporal dynamic semantics of factual counterfactuals, *New Frontiers in Artificial Intelligence, LNAI 3609*, eds. by Akito Sakuai et al., 438-448, Springer.

[Veltman 1996] Veltman, F.: 1996, Defaults in update semantics, *Journal of Philosophical Logic* **25**, 221–261.