Updating beliefs with a consideration about the trstworthiness of sources

Yoshitaka Suzuki

School of Information Science,
Japan Advanced Institute of Science and Technology

2001.2.15

Keywords: belief revision, epistemic relevance, trustworthiness of source, belief base, resolution principle.

What is the epistemic source, for example, the knowledge and the belief? This problem has been discussed in philosophy since antiquity, and it is also researched in artificial intelligence recently, because epistemolgy is useful to research knowledge representations and reasonings, the central concerns in this area. Belief revision, theory change, or belief chage is one of the most important areas about this problem in philosophical logic and AI, and this is the dynamic representation of knowledge and belief, while the other studies are the static representation. In other words, belief revision is the process of incorporating new information into a knowledge base while preserving consistency. The major factor that has made belief revision one of the most active research areas in philosophical logic and AI is a paper in 1985 by Alchourron, Gärdenfors, and Makinson, called AGM. In the AGM framework, there are four epistemic elements, and criteria of rationality. The epistemic elements are epistemic states, epistemic attitudes, epistemic inputs, and epistemic changes. Epistemic states are represented by deductively closed sets of sentences, i.e., $K = Cn(K)$, and they are called belief sets. Three epistemic attitudes about a sentence are that a sentence is accepted, rejected, and indetermined. Operations of epistemic change take the form of either adding or giving up a specified sentences, that is, the form of epistemic inputs. There are three operations of epistemic change. When we contract the belief set K by a sentence ϕ, the result of $K\dot{-}\phi$, called the contraction, is a deductively closed set of sentences that does not contain ϕ. The two operations for the incorporation of new information are expansion and revision. The result of $K + \phi$, the expansion of K by ϕ, is simply $Cn(K \cup \{\phi\})$. The result of $K\dot{+}\phi$ is a deductively closed set of sentences that contains ϕ and is consistent. Hence, in revision but not in expansion, previous beliefs are sacrificed to achieve consistemcy. Criteria of rationality

are the AGM postulates of revision and contraction. AGM proved that the postulates of revision are equivalent to the postulates of contraction by Levi and Harper Identity. Then they defined partial meet contraction and revision, and proved that they satisfy the most of the postulates. Further more, they defined the contraction and revision that satisfy all of the postulates. Both contraction and revision may require deletions of some elements from the original set K of beliefs. there are two major methods for the decision about selecting what to give up. One of the methods is to select among the subsets of K, the other is to select among its elements. One selects by means of a choice function, the other by means of a preference ordering. The original AGM framework has extended in various ways, and Nebel(1992) that is one of these ways used the methods to select among the elements. Nebel wanted to distinguish between the importance or relevance of different sentences in knowledge base, and this idea of assigning different priorities to sentences was formalized by employing a complete preorder with maximal elements, written $\phi \preceq \psi$, called epistemic relevance. Gärdenfors and Makinson(1988) also defined such an preference order, called epistemic entrenchment, that is equivalent to the AGM postulates, but epistemic entrenchment, differ from epistemic relevance, requires the five postulates for the definition, and therefore Nebel's epistemic relevance is more easy to do with than GM's epistemic entrenchement. this preorder is useful as follows, instead of a selection function, the prioritized removal that is defined by this preorder can be used as the method for selecting what to delete. Such a preference relation among subsets, elements, or possible worlds is also used in Doyle(1991), Katsuno and Mendelzon(1992), Meyer et al.(2000), Friedman and Halpern(1997), and etc. but Cantwell(1998) argues that, in spite of many various approaches using preference relation, little work in this area has been done on how such relations come about. He researches the mechanism of how we ought, and ought not, to evaluate the beliefs we have, and he considers that we evaluate information from sources on the basis of their perceived trustworthiness. He then supposes the system of resolving conflicting information with trustworthiness of sources, using the semantics approach that removing possible state without support of sources, but he gives no account on how we come to evaluate sources. In this paper, the method of evaluating sources will be considered as follow; if the information is holded as result of updating knowledge base, the trustworthiness of the source that gives this information is up, otherwise, if the information is given up as result of updating knowledge base, the trustworthiness of the source that gives this information is down. The main purpose in this paper will be to implement this idea, but this implement is difficult, because belief revsion is rather theoretical than implemental. Here the method of implement must be supposed by researching belief revision theory. There are four problems for implementing belief revision system, especially the AGM framework; First, the belief set contains countable infinite elements by deductively closed. Second, The way for judging whether an input is in conflict with knowledge base or not without a logical closure operation, such as Gentzen's natural deduction, must be defined. Third, the computational cost may be very large, that is, the time of execution may increase exponentially. Fourth, if a selection function among the subsets of K is use as the method for selecting what to delete, the way of defining a selection function is unknown. Among the four problems, the last one is already solved by Nebel, as is stated above. The two keys of resolving the

first and second problems are belief base and resolution principle. Belief bases, denoted by the letters A,B,C, are the useful epistemic states, because They are arbitrary sets and don't have to be deductively closed, and therefore they are admitted to be finite sets of sentences. Resolution principle is a fast method for judging whether a sentece is in conflict with a set of sentences, and therefore it is able to understand whether an input is in conflict with knowledge base or not without a logical closure operation. In this paper, instead of Cantwell's semantics approach, Nebel's syntax based approach is adopted as a central idea for implementation. One of the reason for doing so is to be easy to define the way of changing trustworthiness of sources. Thus the implementation of system of updating an epistemic state is tentatively complete, but an experiment explains that the problem of computational cost are still not solved even if Nebel's epistemic relevance is used. For resolving this problem, a method of making the time of execution restricted and stoping this execution can be assumed. How reliable is such an implemented system? An experiment of checking the reliability proves that, if the restriction of time is loose, the system will exactly judge the most reliable source, but if the restriction of time is strict, the system will judge uncertainly. Thus when the restriction is loose, the system is reliable, and this result justify a system with trustworthiness of sources using Nebel's syntax based approach, instead of Cantwell's semantics approach. Then when the exact thought is too sacrificed for shortening the time of execution, the system is looked like a human that failures for putting trust in an unreliable information. After all, the belief revision system can be implemented as a system that updates a knowledge base and is interesting.