ERIC PACUIT, ROHIT PARIKH and EVA COGAN

THE LOGIC OF KNOWLEDGE BASED OBLIGATION*

ABSTRACT. Deontic Logic goes back to Ernst Mally’s 1926 work, Grundgesetze des Sollens: Elemente der Logik des Willens [Mally, E.: 1926, Grundgesetze des Sollens: Elemente der Logik des Willens, Leuschner & Lubensky, Graz], where he presented axioms for the notion ‘p ought to be the case’. Some difficulties were found in Mally’s axioms, and the field has much developed. Logic of Knowledge goes back to Hintikka’s work Knowledge and Belief [Hintikka, J.: 1962, Knowledge and Belief: An Introduction to the Logic of the Two Notions, Cornell University Press] in which he proposed formal logics of knowledge and belief. This field has also developed quite a great deal and is now the subject of the TARK conferences. However, there has been relatively little work combining the two notions of knowledge (belief) with the notion of obligation. (See, however, [Lomuscio, A. and Sergot, M.: 2003, Studia Logica 75 63–92; Moore, R. C.: 1990, In J. F. Allen, J. Hendler and A. Tate (eds.), Readings in Planning, Morgan Kaufmann Publishers, San Mateo, CA]) In this paper we point out that an agent’s obligations are often dependent on what the agent knows, and indeed one cannot reasonably be expected to respond to a problem if one is not aware of its existence. For instance, a doctor cannot be expected to treat a patient unless she is aware of the fact that he is sick, and this creates a secondary obligation on the patient or someone else to inform the doctor of his situation. In other words, many obligations are situation dependent, and only apply in the presence of the relevant information. Thus a case for combining Deontic Logic with the Logic of Knowledge is clear. We introduce the notion of knowledge based obligation and offer an S5, history based Kripke semantics to express this notion, as this semantics enables us to represent how information is transmitted among agents and how knowledge changes over time as a result of communications. We consider both the case of an absolute obligation (although dependent on information) as well as the (defeasible) notion of an obligation which may be over-ridden by more relevant information. For instance a physician who is about to inject a patient with drug $d$ may find out that the patient is allergic to $d$ and that she should use $d'$ instead. Dealing with the second kind of case requires a resort to non-monotonic reasoning and the notion of justified belief which is stronger than plain belief, but weaker than absolute knowledge in that it can be over-ridden. This notion of justified belief also creates a derived notion of default obligation where an agent has, as far as the agent knows, an obligation to do some action $a$. A dramatic application of this notion is our analysis of the Kitty Genovese case where, in 1964, a young woman was stabbed to death while 38 neighbours watched from their windows but did nothing. The reason was not indifference, but none of the neighbours had even a default obligation to act, even though, as a group, they did have an obligation to take some action to protect Kitty.
1. INTRODUCTION

Suppose we are given two functions \( \alpha \) and \( \beta \) over some domain \( D \). Let \( \alpha \leq \beta \) iff \( \forall x \in D, \alpha(x) \leq \beta(x) \), and moreover \( \alpha < \beta \) iff \( \alpha \leq \beta \) and \( \beta \not\leq \alpha \). Suppose now that \( \alpha(d) \ (\beta(d)) \) is the utility value of strategy \( \alpha \ (\beta) \) in some circumstances \( d \). If \( \alpha < \beta \), then we will say that strategy \( \beta \) dominates strategy \( \alpha \). Hence, if some element \( d \) of \( D \) is chosen, and we are offered a choice between \( \alpha(d) \) and \( \beta(d) \) in dollars, we will choose \( \beta(d) \) even if \( d \) is unknown to us. This paradigm comes in useful in two contexts: the decision theoretic context, where \( D \) is the set of possible states of nature and \( \alpha, \beta \) represent payoff functions; and the game theoretic context, where \( D \) represents the (already chosen but unknown to us) choices of the other players, and \( \alpha, \beta \) are possible strategies for us.

Now this comparison between \( \alpha \) and \( \beta \) will not be possible for us if all we are given are the ranges of \( \alpha \) and \( \beta \). For instance if \( \alpha(x) = x^2 \) and \( \beta(x) = x \) over the unit interval \([0,1]\), then it is indeed the case that \( \alpha < \beta \) even though the ranges of the two functions are the same. Moreover, the function \( \gamma(x) = 1-x \) has the same range as \( \beta \), but while we do have \( \alpha < \beta \) we do not have \( \alpha < \gamma \). So the ranges by themselves give us too little information to be able to tell whether \( \alpha < \beta \).

For consider the decision whether to exercise. Suppose some people are rich and some are poor, but all would be better off exercising. However, assume for a moment that it is better to be rich and lazy than to be poor and to exercise. Then the consequences of exercising are \{rich \& exercised, poor \& exercised\} whereas the consequences of being lazy are \{rich \& lazy, poor \& lazy\}. Not all consequences of exercising are better than every consequence of being lazy, even though each individual person, whether rich or poor, is better off exercising. To ask that all consequences of exercising be better than every consequence of being lazy, is too much. So we need to compare situations pairwise, a particular situation with exercising and the “same” situation with laziness. In other words, if choosing between an \( \alpha \) and a \( \beta \), we should choose \( \beta \) if for the specific circumstance we are concerned with, \( \beta \) yields a higher value than \( \alpha \). Choosing intelligently, or responsibly, may require some knowledge about the circumstances.

These considerations have relevance to the situation where the values represent some societal good and we ought to do what is best for society. For knowing what is good may involve knowing some facts.