CMSC 828D Report - Circumscription

Philip Dasler

September, 16, 2012

Introduction

This paper by John McCarthy extends his concept of programs with common sense put forth in his 1959 paper (see references). While not a direct application of his advice-taker, circumscription is a form of non-monotonic reasoning that allows for a program to "jump to certain conclusions" without the need for a fully specified description of a given setting. This, McCarthy posits, is analogous to the way which humans do the majority of their reasoning.

As stated in his earlier work, McCarthy imagines a system in which knowledge about the world is represented by a series of natural language statements. From these, the system is able to infer a "wide class of immediate consequences". In other words, the system does not simply accept the information it is given, but rather uses it to further its understanding of the world. More specifically, the system would reason about the outcomes of particular actions given its understanding of the current state.

This, however, leads to what is known as the "qualification problem". Fully specifying the state of the world so that the system can adequately and correctly infer the outcomes of its actions quickly becomes intractable. Let us consider a simple example: getting to work. In order to get to work, one could take their car. Thus, ownership of a car would be the first piece of information to be specified. But, cars do not work without gasoline (baring, of course, alternative fuels). So now we must specify that there is not only gasoline in the car but also enough to get to work. We must also specify that there is enough oil. And that the roads are not iced over. And that the laws of thermodynamics still behave as they always have. In fact, it seems that no matter how many qualifications are added, still more could be applied. While it may be impossible to capture every single qualification, even capturing a usable subset is likely to require such resources as to make any system of this nature impractical.

Circumscription avoids this problem by, essentially, looking at the world through a more constrained view. It does so by making the assumption that all relevant knowledge is already known. Thus, if you know you must get to work and you know you have a car, then you jump to the conclusion that you can take the car to work. If the level of gasoline in the car was relevant, that information would have been given to you. Thus, the set of relevant information is circumscribed, or restricted, to what is known. While the validity of the conclusions drawn from a circumscribed set of knowledge is dependent on whether or not all of the relevant facts have actually been taken into account, it seems that this form of reasoning does a reasonable job in the face of a non-deterministic world in which not all facts can be known.

Formalism

The concept of conscription can be summed up thusly: a set of objects that can be reasoned, using a set of facts A, to have some property P, are the only objects that satisfy P. More formally, take a sentence A, expressed in first order logic. This sentence will contain some predicate $P(\bar{x})$, where $\bar{x}$ is some tuple. Finally, take $A(\Phi)$ to mean the resulting sentence once all instances of P in A have been replaced by the predicate Φ. Given this, the circumscription of P in $A(P)$ is:

$$A(\Phi) \wedge \forall \bar{x}.(\Phi(\bar{x}) \supset P(\bar{x})) \supset \forall \bar{x}.(P(\bar{x}) \supset \Phi(\bar{x})).$$

This, in essence, asserts the fact that, given the sentence A, certain tuples must satisfy P and these tuples are the *only* tuples to do so. The first conjunct $A(\Phi)$ requires that any arbitrary Φ selected for use in the circumscription still satisfies P. The second conjunct requires that the variables which satisfy Φ are a subset of those that satisfy P (recall that a propositional model can be represented as a set of the propositional variables it assigns to true). The final portion asserts the converse, i.e. the set of variables satisfying P must be a subset of those which satisfy Φ. More importantly, this shows that these two predicates coincide and that the tuple $\bar{x}$ contains all variables relevant to P.

Let us delve into an example as given by McCarthy. Using the famous blocks world problem, we begin with some set of known facts. In the formalism above, this is referred to as the sentence A:

$$isblockA \wedge isblockB \wedge isblockC$$

So, from our sentence, we know that there are three blocks, namely A, B, and C. Now we wish to circumscribe the predice *isblock*. In this case, *isblock* is our predicate P and we wish to say that any knowledge we can derive from A regarding blocks is, in fact, *all* of the relevant knowledge regarding blocks. So, combining our formalism of circumscription with our A and P we get:

$$\Phi(A) \wedge \Phi(B) \wedge \Phi(C) \wedge \forall x.(\Phi(x) \supset isblockx) \supset \forall x.(isblockx \supset \Phi(x)).$$

If we now substitute

$$\Phi(x) \equiv (x = A \vee x = B \vee x = C)$$

into this formula, we can see that the left side is true and can be dropped. Additionally, using the original sentence A removes the next portion as x being equal to A, B, or C necessitates it being a block. This leaves us with:

$$\forall x.(isblockx \supset (x = A \vee x = B \vee x = C)).$$

which states that A, B, and C are the only blocks that exist. Additionally, this example shows that circumscription is nonmonotonic. Recall that the definition of monotonicity involves the addition of information to the system and how it affects the inferences that can be made. That is, in a monotonic system, any inference that can be made can still be made after the addition of further axioms. However, in this system, if we were to add *isblockD* to the sentence A, then we can no longer infer that A, B, and C are the only blocks to exist.

Conclusion

Circumscription allows a system to make assumptions about the world in which it reasons. Rather than being constrained to the knowledge it has, it constrains its view of the world, assuming that all relevant information is known.

Humans reason in much the same way. If one is tasked with going to the store and told that a car is available, the assumption is that the car is functional. If it wasn't, certainly this fact would be relayed as well. We, as humans, reason about the world without complete information all the time. Without the ability to do this, we would suffer from analysis paralysis, stuck in an endless cycle of information gathering, never reaching the point at which a decision can be made. Circumscription allows a system to do the same, by assuming that the information that can be inferred from its knowledge is relevant and complete.

References

J. McCarthy (1980). Circumscription A form of non-monotonic reasoning. *Artificial Intelligence*, 13:2739.

J. McCarthy (1959). Programs with Common Sense. http://www-formal.stanford.edu/jmc/mcc59.html

Circumscription (logic). *Wikipedia.* http://en.wikipedia.org/wiki/Circumscription_(logic)