A Comprehensive Approach to Argumentation

Philip N. Judson and Jonathan Vessey
LHASA Ltd
School of Chemistry
University of Leeds
Leeds LS2 9JT
UK
judson@dircon.co.uk
jonv@chem.leeds.ac.uk

Abstract

A model has been developed which implements the key principles of the Logic of Argumentation – a method of reasoning based on the human process of reaching conclusions by comparing arguments for and against propositions. The model is being used to improve reasoning about the potential toxicity of novel compounds.

Introduction

There is a need to make predictions about toxicities of novel substances, perhaps even before they have been synthesised. Experienced toxicologists are moderately good at doing this, even when the available information is very uncertain. How can computers assist them?

One approach seeks to imitate human experts in knowledge-based expert systems (KBS). There have been some successes in the use of KBS in toxicology. For example, the DEREK system is reported to be quite good at predicting the potential of chemicals as skin sensitizers (Dearden et al. 1997). Current KBS nevertheless fall a long way short of the reasoning capabilities of a human being.

The logic of argumentation (LA) (Krause et al. 1995) was used as the basis of a prototype improved system for predicting toxicity (Fox et al. 1996), but with a restricted model. This paper outlines a more complete model.

LA formalises the way that people reach decisions by identifying the arguments for and against a proposition, and then weighing them against each other. Some arguments relate directly to the decision or proposition of interest. Other arguments may be about the grounds of the direct ones, or about the reliability of the arguments themselves. So the reasoning process can be represented by a tree graph.

Arguments in a chain of reasoning in a complicated domain are rarely conclusive. Indeed, if there is a conclusive argument for or against a proposition, then the matter is decided and all other arguments are irrelevant (save in the case of a conclusive argument in opposition to the first, which is the special case of contradiction).

Reasoning about Toxicity

Three examples illustrate inter-relationships between arguments which a model for reasoning about potential toxicity needs to incorporate.

It is generally thought that chemicals showing carcinogenicity in rats should not automatically be expected to be similarly toxic to other species (including man) if they promote peroxisome proliferation, because this mechanism is peculiar to the rat. Of course, a chemical may be carcinogenic for different reasons, and so it is not valid to state, or to imply, that a compound that is carcinogenic to rats but causes peroxisome proliferation is safe to man – or even that the evidence argues in favour of safety. A reasoning model needs to be able to deal correctly with evidence that undercuts an argument, rather than negates it, like this.

Certain sub-structural features in chemicals are commonly associated with carcinogenicity. In some cases, this connection is not strictly correct, the reality being that positive Ames test results have been linked to the presence of the sub-structural features. The Ames test is an indirect indicator of potential carcinogenicity but this distinction has been lost somewhere along the line between the original researchers and, ultimately, the popular press. A computer reasoning system should construct the overall prediction of carcinogenicity from the basic steps, but it should be able to report its reasoning in full, so that a user is properly informed.

Electrophiles with reactivities within a particular range cause skin sensitization. They must be sufficiently reactive to combine with proteins, leading to the raising of antibodies, but not so reactive that, becoming completely bound to the skin surface, they never reach the site of action. Such reactivity can be assumed with reasonable confidence for a substance that has certain sub-structural features (for example, an alpha-beta unsaturated aldehyde,

which acts as a Michael acceptor). Substances containing the right features do not all provoke sensitization, for a variety of reasons. One is thought to be the failure of a substance to penetrate the skin because of its physico-chemical properties. A computer system needs to be able to modify its predictions about a chemical that meets the criteria for electophilicity, after taking into account physical properties thought to affect skin penetration.

An Outline of the Model

The model we have developed will be described in detail elsewhere (Judson and Vessey, in press), and only an outline is given here. Figure 1 illustrates the main elements of a tree for reasoning in LA.

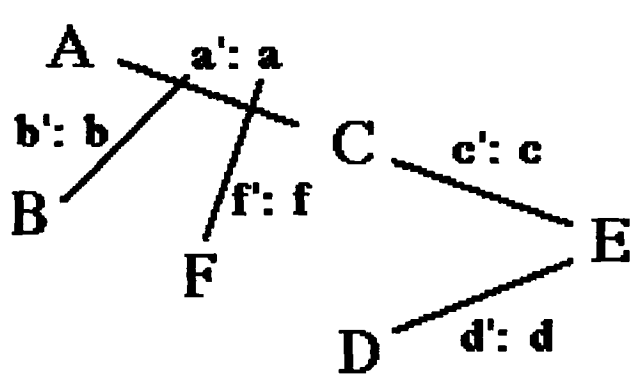

Figure 1 – A simple reasoning tree

The capital letters at the nodes of the graph represent assertions and the arcs represent arguments that connect them. The graph is a directed one, from left to right. So, for example, the tree indicates that the likelihood of C may be deduced from the likelihood of A. The lower case letters represent degrees of persuasiveness of the arguments themselves. So, for example, "a" is how likely C would be if A were proven true (or how unlikely, if the truth of A argues for the falseness of C).

Reasoning along the chain from A to E is like the prediction of carcinogenity from a sub-structural feature associated with positive Ames test results. Step A to C might be "if such-and-such sub-structural feature is present in the molecule, a positive Ames test result is probable". Step C to E might be "if an Ames test result is positive, then carcinogenicity is plausible". A more complete implementation would further sub-divide the chain to include the implicit steps from 'Ames test positive' to "genotoxic" and from "genotoxic" to "carcinogenic".

The persuasiveness of an argument can be subject to other influences, as in the case of predicting toxicity in man from observed toxicity in the rat, where the argument is weakened if peroxisome proliferation is also seen. This is the purpose of arcs leading towards the values of other arcs, such as the one leading from "F" to "a".

In some domains, there may be a case for allowing arguments of the kind "If A is plausible then C is true", which would require the notion of a threshold value for the argument linking A and C. "a'" represents this threshold value. The concept is included for completeness but it will not be discussed further in this paper.

In LA an argument for a proposition is considered to have no direct effect on belief that the proposition is false, and *vice versa*. For example, observed toxicity to an animal may support the proposition that a chemical may be toxic in man: absence of observed toxicity in the animal does not support the proposition that the chemical is non-toxic in man; there is simply a lack of evidence to support the proposition that it is.

This does not mean that negative reasoning is excluded, only that it must be explicit. To assert that the absence of toxicity in laboratory animals makes toxicity in man unlikely, it is necessary to write a rule such as, "If no toxic effects are observed in laboratory animals then it is doubted that the chemical will be toxic in man".

In the implementation of the model which we favour, if there are several arguments in support of a conclusion, the level of belief in the conclusion is equal to the level conferred by the strongest argument. For example, it does not matter how many independent arguments support the view that a conclusion is plausible, it remains plausible. No number of reasons for thinking something is plausible should make it a certainty, to use an extreme illustration.

If the weights of argument for and against a conclusion are equal, belief in the conclusion is classed as equivocal. This is distinguished from the case where there is no evidence either way, which is classed as open. For example, given the information that a chemical shows carcinogenicity in the rat but also causes peroxisome proliferation, it may be appropriate to class the likelihood that the chemical will be carcinogenic in man as open, whereas toxicological concern would be equivocal for a chemical with the right structural features for a skin sensitizer but with physical properties on the borderline for skin penetration.

Given the statement "If A is true then the level of belief in C is a", what should happen if A is less than true? In the model, the level of C becomes whichever is the weaker – the level of belief in A or the value of a. Suppose for illustration, that 'probable' means more

certain than 'plausible' and 'certain' means completely certain. Given the statement "If A is certain then C is probable": C can never be predicted to be more than probable on the basis of evidence about A; if A is probable, then so is C; if A is plausible, then C is also only plausible.

As commented above, if A is proven false, nothing can be said about C on the basis of this evidence. It is not valid to conclude that C is also false, because it may be true for some other reason. More generally, if evidence about the truth of A falls anywhere in the realm of doubt rather than belief, nothing can be concluded about C.

The separation of arguments for and against raises the possibility of contradiction. Theoretically, there might be an argument that proves a conclusion to be true and an independent argument that proves it to be false. Such a situation can arise in practice, too – as the result of erroneous observations, for example, or because of genuine scientific contradiction that has not yet been resolved.

In the implementation of the model which we are using, contradiction dominates among arguments relating directly to the same proposition, unless it is over-ruled by an independent argument that the proposition is true or false: it is usual scientific practice when confronted with apparent contradiction to seek out alternative evidence to resolve the difficulty. Contradiction is not transmitted as contradiction from one level to the next in the reasoning process, but is interpreted as the open state. Given that "if A is true then C is plausible", for example, contradiction in A means that A is either true or false. If it is true then C is plausible. As explained above, if A is false then C is open, which includes the possibility of its being plausible. So, overall, C is open.

The Basic Model

The central components of the model are the methods for aggregating arguments relating to the same proposition (e.g. the combination of degrees of belief in E conferred via the arcs from C and D in figure 1) and for determining the degree of belief conferred by each argument (e.g. the belief transmitted to C from A). They are as follows.

Combining Arguments

The aggregate value for a node, T, is determined by:

$$T = \text{Resolve}[\text{Max}\{\text{For}(C_{a,x}, C_{b,y}, \ldots\},\ \text{Max}\{\text{Against}(C_{a,x}, C_{b,y}, \ldots\}]$$

where: Resolve[] returns the single value which is the resolution of any pair of opposing forces in a resolution matrix (see below); the sets For and Against are the sets of arguments supporting and opposing T, respectively; $C_{a,x}$, $C_{b,y}$, ... are the forces of those arguments; Max{...} returns the member of the set upon which it operates which is highest in an ordered list For or Against, as appropriate (see below).

Reasoning along a Chain

The value transmitted along an arc in the decision tree for the statement, "if a is x_0 then C is x," is determined as follows:

```
if x0 is exclusive then
   if a = x0 then
      C = x
   else
      C = open
   end if
else if x ∈ Subset(¬a) then
   C = open
else if a = Min{a, x0} then
   C = Min{a, x}
else
   C = x
end
```

where: Subset(¬a) is the set of terms For or Against which does not contain value 'a'; Min{a, x} is the term ranking lower in the ordered list, For or Against, to which a and x belong. The term 'exclusive' is explained below.

The Resolution Matrix

For the equation above to operate it is necessary to define a matrix for the resolution of pairs of levels of belief. The matrix which we are using is shown in Table 1.

Ordered Sets For and Against

We are using the following sets:

For	Against
certain	impossible
contradicted	contradicted
probable	improbable
plausible	doubted
equivocal	equivocal
open	open

The operations described above are only valid for terms that can be ordered and resolved unambiguously. For the

purposes of this implementation of the model the terms are defined as follows (note that in some cases the definitions are more restrictive than the meanings of the words in common usage):

certain	there is proof that the proposition is true
probable	there is at least one strong argument that the proposition is true and there are no arguments against it
plausible	the weight of evidence supports the proposition
equivocal	there is an equal weight of evidence for and against the proposition
improbable	there is at least one strong argument that the proposition is false and there are no arguments that it is true
impossible	there is proof that the proposition is false
contradicted	there is apparently proof that the proposition is both true and false
open	there is no evidence that supports or opposes the proposition

Exclusivity

When some terms from the above sets are attached to a given argument or proposition, they imply the terms below them in the list. For example, if something is known to be certain, then the answer to the question "is it plausible?" is "yes". Some terms exclude all others. For example, the state of knowledge about something cannot be open and simultaneously have some other value. The latter terms are described as 'exclusive'.

The members of the sets above are categorised as follows:

Exclusive terms: contradicted, equivocal, open.

Non-exclusive terms: impossible, improbable, doubted, plausible, probable, certain.

Conclusion

This paper outlines a model which we believe can be used for reasoning under uncertainty in any domain. Our current interest is in using it to improve the performance of knowledge based systems for toxicology.

Acknowledgements

We wish particularly to thank John Fox for his vision in promoting the basic concepts of LA and Paul Krause for his advice and comments during the development of the model described here.

References

Dearden, J. C.; Barratt, M. D.; Benigni, R.; Bristol, D. W.; Combes, R. D.; Cronin, M. T. D.; Judson, P. N.; Payne, M. P.; Richard, A. M.; Tichy, M.; Worth, A. P.; and Yourick, J. J. 1997. The Development and Validation of Expert Systems for Predicting Toxicity. *ATLA* 25:223-252.

Judson, P. N. and Vessey, J. A Comprehensive Approach to Argumentation. Submitted to the IMA Journal of Mathematics Applied in Business and Industry, 1999.

Krause, P. J.; Ambler, S. J.; Elvang- Gøransson, M.; and Fox, J. 1995. A Logic of Argumentation for Reasoning Under Uncertainty. *Computational Intelligence* 11(1): 113-121.

Table 1
The resolution matrix

	certain	contradicted	probable	plausible	equivocal	open
impossible	contradicted	impossible	impossible	impossible	impossible	impossible
contradicted	certain	contradicted	contradicted	contradicted	contradicted	contradicted
improbable	certain	contradicted	equivocal	doubted	improbable	improbable
doubted	certain	contradicted	plausible	equivocal	doubted	doubted
equivocal	certain	contradicted	probable	plausible	equivocal	equivocal
open	certain	contradicted	probable	plausible	equivocal	open