Modeling Scientific Discourse

Peter McBurney and Simon Parsons
Department of Computer Science
University of Liverpool
Liverpool L69 7ZF United Kingdom
P.J.McBurney,S.D.Parsons@csc.liv.ac.uk

The Problem Domain

We aim to build intelligent systems which can reason autonomously about the risk of carcinogenicity of chemicals, drawing on whatever theoretical or experimental evidence is available. In earlier work [14], reviewing the literature on methods of carcinogen risk assessment, we catalogued the different types of evidence adduced to support these claims, which may be in the form of: experimental results on tissue cultures, animals or human epidemiological studies; analytical comparisons with known carcinogens; or explication of biomedical causal pathways. Because the research which underpins conclusions in this domain is usually at the leading-edge of the scientific disciplines concerned, evidence from these different sources may be inconsistent or conflicting. Consequently, carcinogen risk assessment usually involves the comparison and resolution of multiple evidences and arguments for and against a particular scientific claim [23, 27].

To represent this domain in an intelligent system, therefore, we first require a philosophical model of scientific enquiry. Which philosophy of science is appropriate for such representation, and why? Next, having adopted such a conceptual model, we will need to formalize it. How can this be achieved? In particular, how may we represent the scientific uncertainty characteristic of knowledge in the carcinogen domain? These questions are the focus of this paper, which outlines our current thinking and approach.

Philosophies of Scientific Discourse

Scientific knowledge accumulates by means of arguments and counter-arguments: e.g. claims for and against particular chemicals being carcinogens are made on the basis of supporting (or denying) evidence. Thus, a philosophy of science which explicitly models the dialectical nature of scientific discourse would seem appropriate. Nicholas Rescher [22], a philosopher of logic and argumentation, claims to have been the first to propose a dialectical framework for the progress of scientific inquiry. Similarly, James Freeman [7], another argumentation theorist, discusses scientific discourse in his study of generic argument structure. Both these approaches are from an argumentation theory perspective rather than from the philosophy of science, and so neither is grounded in, nor engages with, a detailed understanding of actual scientific practice. As a consequence, the frameworks proposed could easily be applied to other, non-scientific, domains.

One novel approach from a philosophy of science perspective is the dialectical model of scientific discourse proposed by Marcello Pera [20]. Pera views the enterprise of science as a three-person dialogue, involving a scientific investigator, Nature and a skeptical scientific community. In this model, the investigator proposes theoretical explanations of scientific phenomena and undertakes scientific experiments to test these. The experiments lead to "replies" from Nature in the form of experimental evidence. However, Nature's responses are not given directly or in a pure form, but are mediated through the third participant, the scientific community, which interprets the evidence, undertakes a debate as to its meaning and implications, and eventually decides in favor or against proposed theoretical explanations. We have adopted Pera's model for our application, and provided Nature with a formal role, but manifested it through those of the other participants.

Although more specific than Rescher's or Freeman's models, Pera's model of modern science as a dialogue game could still be applied to domains which do not share science's success in explaining and predicting natural phenomena. We believe, therefore, that our model of science requires an explanation of the success of science. Some philosophers of science believe this is due to the application of universal principles of assessment of proposed scientific theories, such as the confirmationism of Rudolph Carnap or the falsificationism of Karl Popper [21]. However, we do not share these views, instead believing, with Paul Feyerabend [4], that the standards of assessment used by any scientific community are domain-, context- and time-dependent. This view, that there are neither universal nor objective standards by which scientific theories can be judged, was called "epistemological anarchism" by Imre Lakatos [13]. Moreover, there is a methodological problem with falsificationism in our chosen domain of carcinogenicity. As many have argued (e.g. Hansson [12]), it is not possible to falsify statements of the form *"Chemical $\mathcal{X}$ has carcinogenic effects,"* because one can never eliminate the possibility of very weak effects. For instance, if the effects of a carcinogen at the levels of typical exposure are very small or its actions are long delayed, sample sizes in the millions or billions may be required to have reasonable confidence of identifying the effects in a statistical experiment [25].

Instead of the application of universal principles of assessment of theories, we believe science's success arises in part from applying two normative principles of conduct: firstly, that every theoretical explanation proposed by a scientific investigator is contestable by anyone;[1] and secondly, that every theoretical explanation adopted by a scientific community is defeasible. In other words, all scientific theories, no matter how compelling, are always tentative, being held only until better explanations are found, and anyone may propose these. Note that in saying all conclusions are always defeasible, we are not specifying the manner by which they may be overthrown: defeasibility is thus a more general concept than falsificationism. Contestability distinguishes science from, say, extreme political ideologies, such as Nazism or the Juche philosophy of Kim Il Sung. Defeasibility distinguishes science from, say, traditional religion or creationism. On the other hand, both principles apply to human endeavours commonly thought of as scientific but which may fail criteria such as predictive capability (e.g. paleontology; climatology; macro-economics) or falsifiability (e.g. sociobiology; Freudian psychology).

Formalization and Uncertainty Representation

To build an intelligent system based on these principles, we therefore require a (normative) model of scientific discourse which enables contestation and defeasibility of claims. The approach we have adopted has several components. At the highest level, we are attempting to model a discourse between reasonable, consenting scientists, who accept or reject arguments only on the basis of their relative force. An influential model for debates of this type is the philosophy of Discourse Ethics developed by Jürgen Habermas [9] for debates in ethical and moral domains. We have therefore drawn on Habermas' rules of discourse, which were first fully articulated by Robert Alexy [1].

Next, within this structure, we wish to be able to model dialogues in which different participants variously posit, assert, contest, justify, qualify and retract claims. To represent such activity requires a model of an argument, and we have used Stephen Toulmin's [26] model, within a dialectical framework. To embody our belief in epistemological anarchism, we permit participants to contest any component of a scientific argument: its premises; its rules of inference (Toulmin's "warrants"); its degrees of support (his "modalities"); and its consequences. We believe this is exactly what real scientists do when confronted with new theoretical explanations of natural phenomena [4]. When a scientific claim is thus contested,

[1] At least, by anyone from within the scientific community concerned. While an argument may only be given serious consideration by a scientific community when it arises from a member of that community, there usually are no formal barriers to anyone seeking to join the community. Double-blind reviewing of research papers reinforces this openness.

its proponent may respond, not only by retracting it, but by qualifying it in some way, perhaps reducing its scope of applicability. Arne Naess [18] called this process "precizating", and we seek to enable such responses in the system.

Within Artificial Intelligence (AI), intelligent systems for scientific domains have used argumentation for some time (e.g. [5]). However, these applications have typically involved monolectical rather than dialectical argumentation. More recently, Mandy Haggith [10] developed a dialectical argumentation formalism and applied the resulting system to a carcinogenicity debate. However, the primary focus of her work was on knowledge representation in generic domains of conflict, and so her formalism is not grounded in an explicit philosophy of science. Several philosophers of linguistics and argumentation have articulated models of arguments as dialogue games, e.g. Hamblin [11], Carlson [3] and Walton and Krabbe [28]. These formalisms, which aim to model different types of dialogues in generic domains, were not explicitly intended for encoding in intelligent systems, and they do not in general permit degrees of support for commitments to be expressed. Within AI, the work of Amgoud, Maudet and Parsons [2] is closest in approach to what we are seeking to achieve, although it is focused on generic dialogues and only between two participants.

Within this tradition, we have proposed a formalism which permits participants to make, contest and defend claims, and we call the resulting system a "Risk Agora" [16]. Moreover, by using labels from uncertainty dictionaries, our formalism permits participants to express degrees of belief — in claims, in their supporting evidences, in their modes of inference, and in their consequences. Participants could, for example, accept a scientific claim but label it as, say, *Plausible*, rather than as *Confirmed*. As an example of the formal syntax, a debate participant $\mathcal{P}_i$ could demonstrate her argument $\mathcal{A}(\to \theta)$ supporting a claim θ, an argument to which she was committed with strength D, by making the dialogue move:

$$\textbf{show\_arg}(\mathcal{P}_i : \mathcal{A}(\to \theta, D)).$$

How do we represent Pera's player Nature within this formalism? We have approached this by making Nature an explicit participant, denoted *Nature*, in the scientific dialogue, but one whose interventions in the debate depend on the progress of the debate and on the interventions of the other participants. Thus, for instance, the assessment made by *Nature* of the truth-value (understood in a defeasible sense) of a claim depends, at any one time, on the type and extent of arguments for and against that claim previously articulated by the other participants. As arguments for and against a claim are proposed, contested and defended, *Nature's* assignment of truth-values to the claim may change. In this way, we can explicitly represent scientific uncertainty and the progress of a debate. We are also readily able to summarize the status of the debate, in terms of the relationships between the arguments presented by the debate participants.

In addition to representing the relevant scientific debate, the Risk Agora can potentially assist decision-making in this domain in two other ways. Firstly, the Agora could be used to consider the consequences of statistical inference errors in scientific experiments. Normally, such consequences cannot be quantified and so are ignored in the standard (Neyman-Pearson) procedures used in the bio-medical sciences for hypothesis testing. This has the effect of making the procedures appear to be value-free, when in fact they are not. We have presented a case for using the Risk Agora to design hypothesis tests informed by the qualitative consequences of inference errors in [17].

Secondly, the Agora may assist in the selection of appropriate regulatory options for potential carcinogens. Decision-making here entails assessment of the consequences of different regulatory actions, along with an assignment of values to these consequences. To represent these inside the Agora, we have drawn upon the Communicative Action theory of Habermas [8], which seeks to understand how a group of reasonable people may collaborate to reach a common understanding or decision regarding an issue. In his work, Habermas draws on a typology of speech acts due to John Searle [24], which we have further modified for the specifics of the risk domain [15]. Our work here also draws on our recent work in qualitative decision theory in AI [6, 19].[2]

References

[1] R. Alexy. A theory of practical discourse. In S. Benhabib and F. Dallmayr, editors, *The Communicative Ethics Controversy*, Studies in Contemporary German Social Thought, pages 151–190. MIT Press, Cambridge, MA, USA, 1990. (Published in German 1978).

[2] L. Amgoud, N. Maudet, and S. Parsons. Modelling dialogues using argumentation. In E. Durfee, editor, *4th Intern. Conf. on Multi-Agent Systems*, Boston, MA, USA, 2000. IEEE.

[3] L. Carlson. *Dialogue Games: An Approach to Discourse Analysis*. D. Reidel, Dordrecht, The Netherlands, 1983.

[4] P. Feyerabend. *Against Method*. Verso, London, UK, third edition, 1993. First edition published 1971.

[5] J. Fox, P. Krause, and S. Ambler. Arguments, contradictions and practical reasoning. *Proceedings of ECAI 1992, Vienna, Austria*, 1992.

[6] J. Fox and S. Parsons. Arguing about beliefs and actions. In A. Hunter and S. Parsons, editors, *Applications of Uncertainty Formalisms*, LNAI 1455, pages 266–302. Springer Verlag, Berlin, Germany, 1998.

[7] J. B. Freeman. *Dialectics and the Macrostructure of Arguments: A Theory of Argument Structure*. Foris, Berlin, Germany, 1991.

[2] We are grateful for partial funding from the British EPSRC under grant GR/L84117 and a PhD studentship. We also thank Trevor Bench-Capon, Mark Colyvan, Susan Haack, Adrian Heathcote, William Rehg and the anonymous referees for their comments.

[8] J. Habermas. *The Theory of Communicative Action: Volume 1: Reason and the Rationalization of Society*. Heinemann, London, UK, 1984. (Published in German 1981).

[9] J. Habermas. *Moral Consciousness and Communicative Action*. MIT Press, Cambridge, MA, USA, 1991. (Published in German 1983).

[10] M. Haggith. *A Meta-level Argumentation Framework for Representing and Reasoning about Disagreement*. PhD thesis, University of Edinburgh, UK, 1996.

[11] C. L. Hamblin. Mathematical models of dialogue. *Theoria*, 37:130–155, 1971.

[12] S. O. Hansson. What is philosophy of risk? *Theoria*, 62(1-2):169–186, 1996.

[13] I. Lakatos and P. Feyerabend. *For and Against Method*. University of Chicago Press, Chicago, IL, USA, 1999.

[14] P. McBurney and S. Parsons. Truth or consequences: using argumentation to reason about risk. In *Symposium on Practical Reasoning*, London, UK, 1999. British Psych. Soc.

[15] P. McBurney and S. Parsons. Intelligent systems to support deliberative democracy in environmental regulation. In D. Peterson, editor, *AISB Symposium on AI and Legal Reasoning*, Birmingham, UK, 2000. AISB.

[16] P. McBurney and S. Parsons. Risk agoras: Dialectical argumentation for scientific reasoning. In C. Boutilier and M. Goldszmidt, editors, *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence (UAI-2000)*, Stanford, CA, USA, 2000. UAI.

[17] P. McBurney and S. Parsons. Risk Agoras: Using dialectical argumentation to debate risk. *Risk Management*, 2000 (In press).

[18] A. Naess. *Communication and Argument: Elements of Applied Semantics*. Allen and Unwin, London, UK, 1966. (Published in Norwegian 1947).

[19] S. Parsons and S. Green. Argumentation and qualitative decision making. In A. Hunter and S. Parsons, editors, *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, LNAI 1638, pages 328–339. Springer, Berlin, Germany, 1999.

[20] M. Pera. *The Discourses of Science*. University of Chicago Press, Chicago, IL, USA, 1994.

[21] K. R. Popper. *The Logic of Scientific Enquiry*. Hutchinson, London, UK, 1959.

[22] N. Rescher. *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*. State University of New York Press, Albany, NY, USA, 1977.

[23] L. R. Rhomberg. A survey of methods for chemical health risk assessment among Federal regulatory agencies. *Human and Ecological Risk Assessment*, 3 (6):1029–1196, 1997.

[24] J. Searle. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge University Press, Cambridge UK, 1979.

[25] M. E. Shere. The myth of meaningful environmental risk assessment. *Harvard Environmental Law Review*, 19 (2):409–492, 1995.

[26] S. E. Toulmin. *The Uses of Argument*. Cambridge University Press, Cambridge, UK, 1958.

[27] U.S.A. Environmental Protection Agency. Proposed guidelines for carcinogen risk assessment. *U.S. Federal Register*, 61:17960–18011, 23 April 1996.

[28] D. N. Walton and E. C. W. Krabbe. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press, Albany, NY, USA, 1995.