ARAUCARIA: MARKING UP ARGUMENT

1. Introduction and Motivation

Argumentation theory aims to analyse, describe, and evaluate real-world, natural language arguments, and occurs as a topic in many undergraduate syllabi, where it aims to teach students both to think critically about the arguments of others, and to create better, more measured arguments of their own (see (van Eemeren *et al*, 1996) for an overview of the field). One of the key tools available to the discipline, in both its academic and pedagogic branches, is diagramming. The claims and their associated reasons within a given argument are identified, and the relationships between them drawn up as trees. This diagram then serves as a basis for criticism and reflection.

Computer software might well be anticipated to be highly suited to the task of visualisation, and particularly to the sort of diagrams used in argumentation theory. Similarly, it is also well suited to the task of aiding an analyst in constructing the diagrammatic representation of an argument. And yet, there are very few computer systems which support argument diagramming for the student, and none at all which support the diversity and sophistication of analyses formed within the research community. It is this dearth of computer support for a labor intensive but crucial activity, which is addressed in this paper.

2. Background

The development of informal logic and argumentation theory within philosophy has represented a backlash against post-Fregean formal logic, which though immensely powerful and widely applicable, is a poor choice for representing and characterising natural - i.e. real-world - language and argument, despite its Aristotelian heritage aimed at just that. The inception of informal logic marked by Toulmin (1958), Perelman and Olbrechts-Tyteca (1969), and others saw a return to an empirically driven logic.

Within argumentation theory, systems of diagramming argument have played an important practical role in two distinct areas. The first is in pedagogy: employing diagrams in support of the teaching of critical thinking skills. Though opinion is divided as to the degree to which diagramming is useful for all students (see, for example, (Argthry, 2001)) there is clear evidence that the technique is of great benefit for some (van Gelder & Rizzo, 2001). These results concur with more general results in the psychology of reasoning (see (Rips, 1994: 347-349) for an introductory discussion of this issue).

The driving force provided by the need to teach critical thinking, and the rise in utility and subsequent popularity of computer assisted learning packages, has led to the appearance of a number of software systems for argument diagramming which are intended for pedagogical use. One of the most sophisticated and polished examples is the Reason!Able system (van Gelder & Rizzo, 2001). This system aids students - particularly those learning informal reasoning skills at an introductory level in schools and universities (and, they report, even kindergarten) - in constructing and analysing argument maps. These maps employ arrows and colours to indicate support and rebuttal relationships, and are manipulated through a straightforward interface that is appealing to children. The preliminary results reported in (van Gelder & Rizzo, 2001) suggest that the software has, at the very least, the potential to substantially improve students' critical thinking skills.

The second role of diagramming is in the construction and implementation of theories of argument evaluation within the research community. One of the earliest methods, now the textbook favourite, is that proposed by Beardsley (1950), and enhanced with nomenclature by Thomas (1986). More recently, inadequacies and problems with this *standard treatment* have been identified, leading several authors to propose alternatives, e.g. Freeman (1991) who extends the standard treatment to deal with structures described by Toulmin (1958); Reed (1999) who explores an alternative method of handling linked structures; and Wilson (1986) who emphasises the evaluative aspect of argument analysis by including it explicitly in the diagram.

Although there has not to date been software specifically designed to support research into argumentation and diagramming, there have been a few systems which impact upon that research. Foremost amongst these is the ambitious and far-reaching *Archelogos* project under development at the University of Edinburgh (Scaltsas, 1997), which aims to analyse and mark-up substantial portions of the argumentation in the oeuvres of both Plato and Aristotle, and provide an interface that allows online navigation of the structure of the reasoning in the works. The *Archelogos* project does not, however, focus upon diagramming the structure that is produced through analysis. Work in linguistics, and in particular in pragmatics, aiming to analyse interclausal relations, does employ software tools to build diagrammatic analyses of textual structure. RSTtool (O'Donnell, 1997) is a good example of such software, and it has been argued (Mann and Thompson, 1988) that the approach can be applied to argumentative text (just as it can to any other genre). These linguistic research projects make no use, however, of the rich analytical structures and techniques of argumentation theory.

Finally, argumentation itself has found many applications within computer science, and various branches of artificial intelligence in particular. A review of many of these systems can be found in (Reed, 1997), whilst a more recent analysis of interdisciplinary work between argumentation and each of multi-agent systems, legal reasoning decision support, computational linguistics, and contextual reasoning, can be found in (Reed and Norman, 2003).

The focus here, however, is squarely upon software to support both teaching and research in argumentation theory. To the authors' knowledge, there is currently no system which provides such support, and it is this gap that the *Araucaria* system fills.

3. Araucaria

The Araucaria system does not attempt to tackle fundamental problems in the diagramming process. As with other methods of analysing textual structure - such as Rhetorical Structure Theory, RST (Mann and Thompson, 1988), for example - any given analysis is potentially disputable. RST offers a means of specifying the relation that holds between spans of text - though both the judgements concerning the delimitation of text structure phrases, and the identification of relationships between those phrases, can be challenged. Mann and Thompson suggest that in marking up the rhetorical structure of a text, the analyst makes *plausibility judgements* (rather than absolute analytical decisions) and that there can be more than one reasonable analysis. The assumptions behind Araucaria follow the same pattern: a single text might be analysed in several different ways, depending upon a variety of analytical choices, and upon the aims of the analyst.

Again by analogy to RST, there is also freedom in analytic resources. Mann and Thompson emphasise that the set of relations they put forward is simply one possible set that has been found to have utility in the analysis of a particular corpus. They claim neither exhaustiveness nor accuracy of their proposed set, instead describing the process by which researchers can produce their own sets of relations. A similar solution is adopted in the provision of *schemesets* of argumentation schemes. Many scholars and teachers of critical thinking and related fields find that argumentation schemes are a useful tool for describing the relationships between argument components. Determining a single, exhaustive, consistent set of schemes has proved difficult - though existing sets such as (Grennan, 1996), (Katzav and Reed, 2003) (Kienpointner, 1992), (Walton, 1996) are nevertheless rich and extensive. The importance of argumentation schemes is also growing within various computational applications of argument (Reed and Walton, 2001), so one aim in developing the Araucaria software was to ensure that argumentation schemes were coherently integrated. The choice of which - if any - argument set to use is left to the user, with the standard distribution including not only schemeset definitions corresponding to the Grennan, Kienpointner and Walton lists mentioned, but also software to design custom schemesets in a straightforward manner.

The emphasis upon comparison with Rhetorical Structure Theory is quite deliberate. By accepting the diversity not only of language, but also of the interpretation and analysis of language, RST has become a powerful and widely used tool in discourse analysis and computational linguistics, and has played a key role in making common resources available to the research community. By equipping argument analysis tools with a similar flexibility and tolerance of analytic diversity, the rich variety of approaches in teaching, learning, and research can be preserved whilst at once providing a common interlingua and environment for carrying out those activities.

As part of the commitment to supporting diversity, Araucaria has been developed in Java, to support execution on many platforms. The software has been tested under various versions of Microsoft Windows, Solaris, Linux and MacOS.

System Overview

The main Araucaria window is shown in Figure 1, below. The system can load either a text file or an existing, marked up, argument. In either case, the left-hand pane shows the original text of an argument. Selecting part of this text and dropping the selection on to the right-hand, diagram, pane creates a node which corresponds to that text. There are then a number of actions which structure and combine the nodes in the diagram:

- Dragging a line from one node to another creates a support relationship, indicated by an arrow from premise to conclusion (supporter to supported)
- Multiple (or 'convergent') lines of support can be selected and 'linked' together (and then subsequently 'un-linked' back to convergent support - Araucaria adopts the terminology of Freeman (1991), *inter alia*)
- Selecting either multiple lines of support or multiple nodes allows the selection of an argumentation scheme to be associated with those nodes. Schemes are identified by coloured areas on the graph, and can be overlapping (that is, a single text span can play a role in several argumentation schemes - typically as a premise in one and a conclusion in another)

In addition, the menus also support a range of further functionality:

- Addition of missing premises to the diagram
- Creation, modification, loading and saving of argumentation scheme sets
- Inclusion of refutations
- Identification of owners, giving an ability to represent multi-party arguments

File manipulation supports the saving not only of the marked up argument using XML, but also of the diagram itself as a JPEG image for incorporation into documents and online material. Finally, as opinion seems divided in both the research and pedagogic communities, the entire diagram can be inverted at any time with a single key-press.

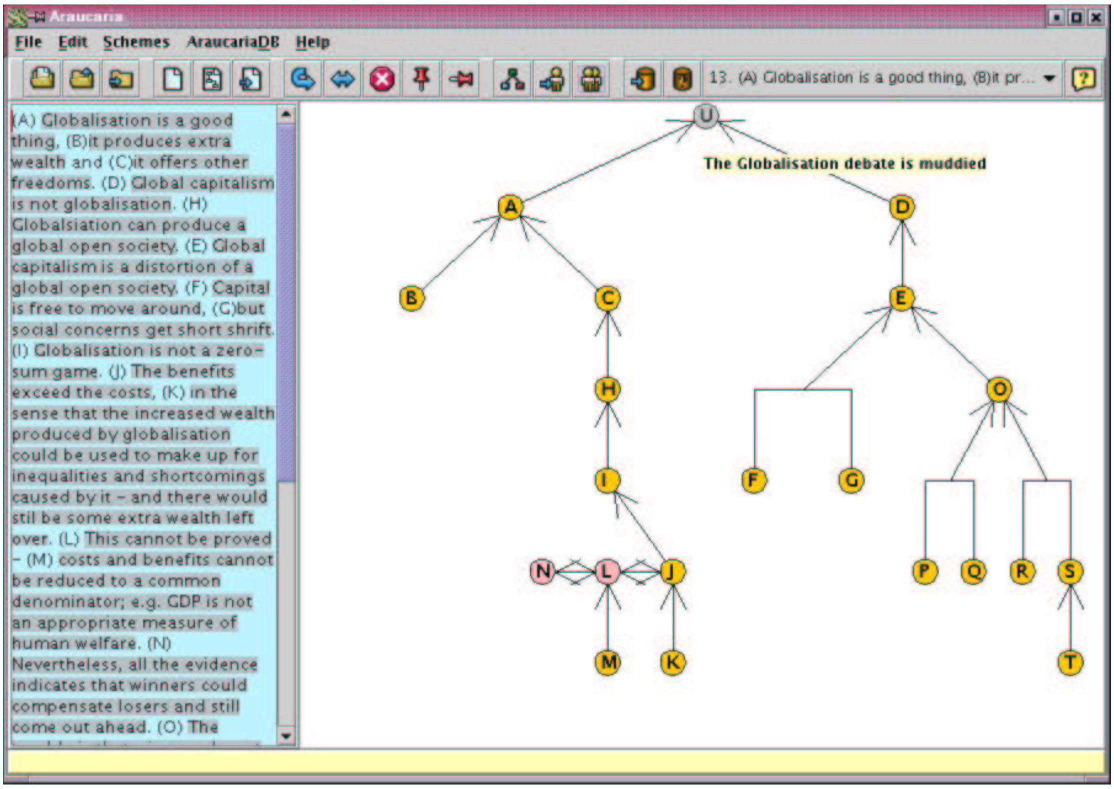

Figure 1. Araucaria

4. AML

The environment provided by Araucaria is one suited to analysis. That is, it is assumed that a sample text is available and that this text is to be analysed to produce a diagram. It is important during this process that links between the original text and the corresponding components of the diagram be maintained: deletions of components of the diagram, for example, rely upon these links. Further, once the analysis is complete, it is useful to save not just the text and diagram, but also the relationship between them, to allow future modification and manipulation.

These factors suggest that marking up the original text is an appropriate approach. By adding to the original text, *tags* that indicate the evolving structure, the relationship between text and diagram is preserved. The *argument markup language* (AML) defines a set of tags that indicate delimitation of argument components (loosely, propositions), tags that indicate support relationships between those components, and tags that indicate the extent of instances of argumentation schemes.

Both Araucaria, and the markup language in which analyses are saved, exploit the typical tree structure of argumentation. This means that the markup language can be defined quite succinctly, by characterising an argument recursively as a proposition supported by one or more arguments. Implementation of the markup language employs XML (eXtensible Markup Language) which carries with it a range of benefits. Firstly, XML is a well-defined, well-understood, widely used industrial standard. This means that there are a range of standard tools which can be employed to manipulate the data: one simple example is a sample application available with many XML parsing suites. The application provides a conventional tree view (such as is employed by Microsoft Windows Explorer) that, in the context of the files produced by Araucaria, allows a conclusion to be expanded to show its (immediate) supporting premises, and each of those premises to be expanded to show their supports, and so on. There are similarly tools for creating summaries, diagrams, for verifying content, etc.

Secondly, as a generic data representation language XML files are also easily translatable into other formats through the application of *stylesheets*. A good example of the possible uses of stylesheets is in the creation of tailored HTML web files. Thus arguments can be automatically summarised or made navigable for online provision. The discussion returns to various such applications in sections 5, below.

Thirdly, XML has recently been recognised as having the potential to play a significant role in corpus resources (Ide, 2000). Since this is one of the areas of application planned for Araucaria and AML as described in section 5, the adoption of XML leads to several advantages in the development of, publication of, and access to, corpus resources.

Fourthly, the acceptance of XML as a de facto industry standard facilitates data sharing: with a single common format or interlingua, separate applications can share data in the tasks of input, manipulation and output. With the structure of a particular XML markup language defined independently in a document type definition (DTD) file, the definition of AML can quickly and simply be made open to the community through the publication of its DTD .

Finally, the definition of AML in its DTD is independent of the applications in which it is employed. This independence facilitates maintenance and development, whilst permitting backward compatibility. AML is defined in `argument.dtd`, a full specification of the components from which arguments - and argument diagrams - can be constructed[1].

The Araucaria software is thus uncoupled from the AML; the latter can evolve (monotonically) without requiring changes to the former. Furthermore, the stability of the AML also facilitates the development of a suite of related applications.

5. Application Areas

There is a very wide range of potential applications; in this section, the focus is squarely upon the current and potential applications of the currently implemented system functionality.

The foremost application domain is pedagogy. Teaching critical thinking skills, particularly in North America, forms an important part of the curriculum in providing generic, transferable skills. Syllabi for the topic, such as those provided in popular text books such as (Johnson & Blair, 1993; Govier, 1997; Groarke *et al.*, 1997) typically introduce some method for diagramming arguments fairly early on, to provide students with the practical scaffolding around which to erect a battery of analytic techniques. Though the various techniques may differ somewhat, and the presentation of them differ significantly in these works, the diagramming tools are substantially the same. Thus the diagramming itself is uncoupled from the subsequent presentation of critical thinking skills and techniques, which suggests that a‘theory-neutral’ software tool such as Araucaria might be successfully employed as a component of teaching support in a broad range of argumentation and critical thinking courses. The current version of the software is being trialled on undergraduate courses at

1 The file `argument.dtd` can be downloaded from the project homepage at
`http://www.computing.dundee.ac.uk/staff/creed/research/araucaria`

several Canadian and US universities in the fall of 2002.

With domain information structured as arguments in AML, there is also a rich potential for supporting the teaching of other topics. Thus, for example, a small corpus of AML arguments, perhaps constructed using Araucaria, could capture the material for part of an introductory paleontology course, giving arguments for and against conclusions to be drawn from various aspects of the fossil record[2]. Student interaction with this resource could then form part of a computer assisted learning environment such as those described in (Jackson, 1998). Employing applications currently under development, students could review arguments and summaries of them, engage in dialogic exchanges, extend the existing arguments with their own additions, and so on.

In a similar vein, such applications could also have a role to play outside the classroom, in providing one resource for topics in the Public Understanding of Science and Technology (PUST). With online provision of the same tools (Araucaria, dialogic interaction, dynamic generation of summaries, etc.) and a set of AML resources in topical areas such as the genetic modification of food, conflicting viewpoints could be presented to the public in a coherent and measured way. It would be simple to provide for public interaction with the arguments, supporting the contribution of new arguments to the online database. Using arguments to structure online debate has been found to be a good means of involving people in public policy decision making processes (Gordon & Karacapilidis, 1997), and it might be expected that similar advantages might accrue in PUST.

Araucaria and the underlying data representation format also has the potential to serve the academic community in several respects. First, from a practical point of view, the output of Araucaria, both graphical and textual, simplifies the task of preparing material for dissemination. Secondly, and equally practically, having a common format in which to express the structure of an argument simplifies the task of exchanging and dissecting analyses. (Argthry, 2001) has many good examples of academic discussion on structural analyses of problematic examples; each suggestion and counter-suggestion is phrased in idiosyncratic and lengthy analyses which increase the chances of misunderstanding and error. A common language may not make the analyses themselves any easier, but it will at least make the subsequent exchange and discussion of those analyses more open, and less prone to confusion.

2 Paleontology as a discipline offers particularly rich examples of texts in which dialectical structure and chains of argumentation are extremely clear. It is for this reason that many introductory texts employ a chronological basis for exposition, following the various turns of the academic dialogue. See, e.g. (Edey & Johanson, 1990) and chapters 15-16 of (Wilford, 1985)

Much more substantial than these, however, is the provision of a corpus of argumentation, analysed and marked up in AML, and made available online. Work is starting at Dundee to construct such a corpus, the contents of which will be accessible from the WWW, and from Aracuaria. Because the data is stored in a highly structured form using AML, sophisticated access and manipulation becomes possible: a visitor might search for arguments with a particular structure, or for examples of a particular argumentation scheme, or for arguments in a particular domain, or for arguments with a particular degree of complexity, and so on. Furthermore, if Araucaria is used by both academics and students in argumentation to mark up new examples of arguments, then these analyses can be submitted to the corpus to extend the resource for others.

6. Conclusions

The Araucaria system performs a range of functions which are unique, and the software has the potential to play a significant role in both academic and educational domains. Perhaps the most similar software is van Gelder's (2001) Reason!Able system. Like Reason!Able, Araucaria employs a tree structure for mapping out the relationships between components in an argument, and allows the user to manipulate that structure. Unlike Reason!Able, however, Araucaria is driven primarily by research concerns rather than educational concerns, and although pedagogy is a significant application area for Araucaria, it is not the only such area. As a result, the current version utilises, for example, recent research on argumentation schemes to provide support for analyses based upon such schemes. In addition, Araucaria also starts with the assumption that the task at hand is one of analysis of existing argument, rather than the construction of a new argument; for Reason!Able the focus is squarely upon argument synthesis. Finally, Araucaria differs fundamentally from Reason!Able and all other argumentation software in its provision of AML, an open standard for argument description defined in XML, which has the potential to have a pervasive effect in both teaching and research.

Though many of the applications described in section 5 are currently under development, they are described here to show the role that the implemented Araucaria system and its underlying data representation format, AML, will play in a variety of domains. Though Araucaria on its own represents a significant tool for those working in argumentation, when coupled with the applications in the domains suggested, it has the potential not only to play a key role in the development of a range of systems of real utility in academic, pedagogical, and public arenas, but also to support and encourage the further development of aspects of argumentation theory and the application of that theory in computer systems.

Araucaria is free, open-source software, released under the GNU General Public License. The software can be downloaded from the project homepage,
`http://www.computing.dundee.ac.uk/staff/creed/araucaria/`

Alternatively, a distribution CD can be obtained by sending an email to
`araucaria@computing.dundee.ac.uk`

References

Argthry (2001) Discussions on the ARGTHRY list, particularly the archives for July 2001 (available through the process described at `http://www.yorku.ca/gilbert/argthry/`)

Beardsley (1950) *Practical Logic*. Prentice Hall.

Bunke, H & Messmer, BT (1997) "Recent advances in graph matching", *International Journal of Pattern Recognition and Artificial Intelligence*, **11** (1): 169-203.

Edey, M.A. & D.C. Johanson (1990) *Blueprints: Solving the Mystery of Evolution*, Oxford University Press.

van Eemeren, F.H., Grootendorst, R. & Snoeck-Henkemanns, F. (1996) *Fundamentals of Argumentation Theory*, LEA.

van Eemeren, F.H., Grootendorst, R. & Snoeck Henkemanns, F. (1993) *Reconstructing Argumentative Discourse*, LEA.

Freeman, J.B. (1991) *Dialectics and the MarcoStructure of Arguments*, Foris.

van Gelder & A. Rizzo (2001) “Reason!Able Across the Curriculum” in *Is IT an Odyssey in Learning? Proceedings of the 2001 Conference of the Computing in Education Group of Victoria.*

Gordon, T.F. & Karacapilidis, N. (1997) “The Zeno Argumentation Framework” in *Proceedings of the 6th International Conference on AI and Law (ICAIL’97).*

Govier, T. (1997) *A Practical Study of Argument*, Wadsworth.

Grennan, W. (1996) *Informal Logic: Issues and Techniques*, McGill-Queens University Press.

Groarke, L., C.W. Tindale & L. Fisher (1997) *Good Reasoning Matters*, 2nd Edition, Oxford University Press.

Ide, N. (2000) “The XML Framework and Its Implications for Corpus Access and Use” in *Proceedings of LREC’2000 MMLR Workshop*, Athens.

Jackson, S. (1998) “Beyond Web Course Automation: Designing Online Dialogue” in *Working Notes of NAU/Web’98*

Johnson, R. & J.A. Blair (1993) *Logical Self Defense*, McGraw Hill.

Katzav, J. & Reed, C.A. (2003, forthcoming) “On argument schemes and a natural classification of argument”

Kienpointner, M. (1992) “How to classify arguments” in F.H.van Eemeren, R. Grootendorst, J.A. Blair & C.A. Willard, *Argumentation Illuminated.*

Mann, W.C. & S.A. Thompson (1988) “Rhetorical Structure Theory: A Theory of Text

Organisation" *Text* **8** (3) 243-281.

O'Donnell, M. (1997) "RST-Tool: AN RST Analysis Tool", in W. Hoeppner (ed.) *Proceedings of the 6th European Workshop on Natural Language Generation (EWNLG'97),* Duisburg.

Perelman, Ch. and L. Olbrechts-Tyteca (1969) *The New Rhetoric*, Notre Dame Press.

Reed, C.A. (1997) "Representing and Applying Knowledge for Argumentation in a Social Context" *AI & Society* **11**:138-154.

Reed, C.A. (1999) "Building Monologue" in *Proceedings of the 3rd OSSA Conference (OSSA'99),* St. Catharines.

Reed, C.A. & Norman, T.J. (2003, to appear) *Argument and Computation*, Kluwer

Reed, C.A. & Walton, D.N. (2001) "Applications of Argumentation Schemes" in *Proceedings of the 4th OSSA Conference (OSSA'2001)*, Windsor.

Rips, L.J. (1994) *The Psychology of Proof*, MIT Press

Scaltsas, T. (1997) "Representation of Philosophical Argument" in T. Bynum and J. Moor (eds) *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell.

Thomas, S.N. (1986) *Practical Reasoning in Natural Language*, 3rd Edition, Prentice Hall.

Toulmin, S.E. (1958) *The Uses of Argument*, Cambridge University Press.

Walton, D.N. (1996) *Argumentation Schemes for Presumptive Reasoning*, LEA.

Wilford, J.N. (1985) *The Riddle of the Dinosaur*, Faber and Faber.

Wilson (1986) *The Anatomy of Argument*, University Press of America.