Visualizing Carneades Argument Graphs

Thomas F. Gordon
Fraunhofer FOKUS, Berlin, Germany
thomas.gordon@fokus.fraunhofer.de

July 2, 2007

Abstract

Carneades is a computational model of argument, based on the state-of-the-art of argumentation theory in philosophy. This article presents a diagramming method for Carneades, similar to Wigmore charts, and illustrates how to map legal evidence using this method. With suitable computer support, in the form of a special purpose argument diagram editor, users need not understand the mathematics of the computational model to make use of its features. Compared to a generic diagram editor, or even special purpose argument diagramming tools based only on informal models of argument, an argument diagramming tool based on the Carneades computational model of argument has the advantage of being able to inform users, in an intuitively comprehensible way, whether or not a claim satisfies a proof standard, given the evidence and other arguments which have been put forward by the parties. The presentation is entirely informal. No prior expertise in argumentation theory, mathematics or computer science is presumed.

1 Introduction

Carneades is a formal, mathematical model of argument structure and evaluation, based on the state-of-the-art of argumentation theory in philosophy [Walton, 2006], which applies proof standards to determine the acceptability of statements on an issue-by-issue basis [Gordon et al., 2007].[1] Such formal models are essential for developing computer systems which are capable of providing more support for argumentation tasks than is offered by generic diagram editors, such as Microsoft Visio, or even special purpose argument diagramming tools based only on semiformal models of argument, such as AusThink's Rationale or Araucaria [Rowe et al., 2003]. In particular, the Carneades model of argument enables software to inform users whether or not a claim satisfies a proof standard, given the evidence and other arguments which have been put forward by the parties and represented in the diagram. Users need not understand the underlying mathematics in order to use and benefit from such tools, anymore than one needs

[1] I would like to acknowledge Terrance Anderson, the anonymous reviewers and, in particular, Doug Walton, for their helpful comments on draft versions of this article.

to understand the formal specification of word processors when writing documents.[2]

Here we present a diagramming method for Carneades, similar to Wigmore charts, and illustrates how to map legal evidence using this method. No prior expertise in argumentation theory, mathematics or computer science is presumed. The presentation is entirely informal, at the level of a user manual for a diagramming tool for lawyers interested in mapping evidence.

The rest of this article is organized as follows. The next section presents our method for diagraming arguments. The Carneades model of argument will be explained and illustrated, informally, using example diagrams. The following section compares our method with Wigmore charts, using a reconstruction in Carneades of the Umilian case used by Wigmore to illustrate his charting method. The final section summarizes the main results and discusses possible future work.

2 A New Argument Diagramming Method

Figure 1 presents a first, simple example of our argument diagramming method. It is a reconstruction of the 'Nixon diamond' example often used in the field of Artificial Intelligence. This issue is whether or not Richard Nixon is (was) a dove, rather than a hawk. Nixon was a Quaker. This is offered as evidence, in argument a1, of him being a dove, since Quaker's generally are pacifists. On the other hand, Nixon was a Republican, who have a reputation for being hawks, not only during Nixon's lifetime. Argument a2 in the figure expresses this counterargument.

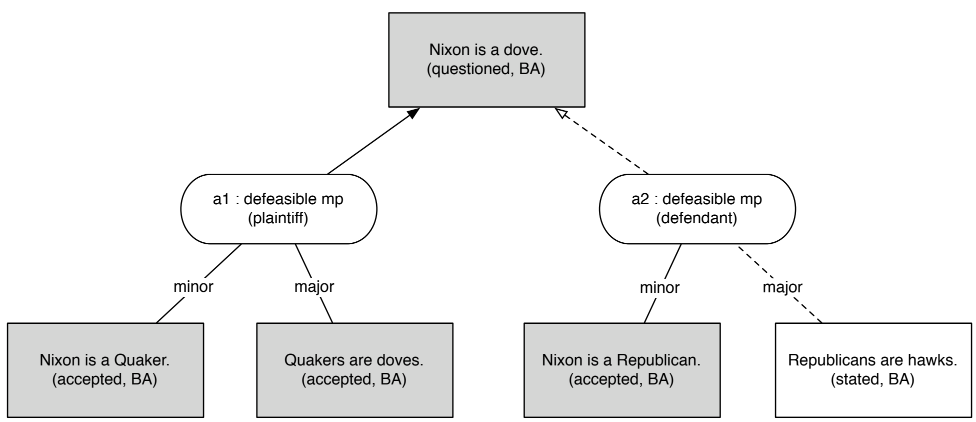

Figure 1: Nixon Diamond

Arguments are displayed as rounded boxes. Information about the argument is shown inside the box. Consider argument a1, on the left. The box includes the name of the argument (a1) and, in parentheses, the name of the party who put forward the argument (plaintiff). Arguments can be instances of general patterns of argument, called

[2]For example, the specification of Microsoft's XML format for word processing documents is several thousand pages long.

‘argumentation schemes’ [Walton, 2006]. The user can label an argument with the name of the argumentation scheme applied. In the example, argument a1 is labeled as being an instance of an argumentation scheme called defeasible modus ponens, abbreviated to ‘defeasible mp’ in the diagram. Where this information comes from depends on the user’s role and task. If the user is an analyst trying to ‘reconstruct’ arguments from some text, determining which argumentation scheme was applied requires the text to be interpreted in light of knowledge of some set of argumentation schemes. Argumentation schemes may be domain-dependent, so particular schemes may apply to the problem at hand. In the legal context, schemes may be derived from the applicable substantive law of the jurisdiction [Gordon, 2005]. If the user is an advocate using the diagramming software to help prepare his case, the software might provide a set of argumentation schemes to choose from, in which case the label of the selected argumentation scheme would be displayed.

An argument links a set of statements, called premises, to another statement, called the conclusion. In diagrams, statements are displayed in rectangles. In the example, “Nixon is a dove.” is the conclusion of both arguments, a1 and a2. The label of statement nodes shows the content of the statement (or a concise identifier) and, in parentheses, the procedural status of the statement (stated, questioned, accepted or rejected). When a statement is first used in a procedure, it is ‘stated’. If the statement is later called into question, it becomes ‘questioned’. After a decision has been taken, presumably after some period of discussion in which arguments pro and con are exchanged, the statement is either ‘accepted’ or ‘rejected’. Also shown in the label of a statement is the name or identifier of the proof standard which applies to the statement in the procedure. Computational models of three such proof standards are presented in [Gordon et al., 2007], called ‘scintilla of evidence’ (SE), ‘best argument’ (BA) and ‘dialectical validity’ (DV).[3] In the example, for simplicity we have assumed that the BA standard applies to all statements.

The links between an argument and its premises and its conclusion also carry information. The link from a premise to an argument can be labeled with the role of the premise in the argumentation scheme applied. In the example, both arguments apply the defeasible modus ponens argumentation scheme, which has major and minor premises, as shown in the diagram.

As for the link from an argument to its conclusion: pro and con arguments are distinguished using different kinds of arrowheads. A solid arrowhead is used for pro arguments, as in argument a1; an open arrowhead is used to indicate con arguments, as in argument a2.

A statement which satisfies its proof standard, given the arguments in the diagram, is said to be ‘acceptable’. In this diagraming method, the rectangles of statements which are acceptable, as well as those which have been accepted, are filled with a gray background. In the example, the main issue, “Nixon is a dove.” is shown in gray, indicating that the evidence represented by the given arguments satisfies the ‘best argument’ (BA) proof standard. The BA standard is satisfied only if the statement is supported some defensible pro argument with priority over all defensible con arguments.

[3]The legal proof standards of ‘preponderance of the evidence’, ‘beyond a reasonable’ and ‘clear and convincing evidence’, among others, have yet to be formally modeled.

In Carneades, an argument is defensible only if all of its premises 'hold'. Whether or not a premise holds depends on its type (ordinary, exception or assumption), status (stated, questioned, accepted, or rejected) and the acceptability of its statement. See [Gordon et al., 2007] for further details. In the example, "Nixon is a dove." is acceptable, since it is supported by a defensible pro argument, a1, and its only con argument, a2, is not currently defensible. The con argument is not currently defensible because its major premise, 'Republicans are hawks.' has not been supported by any evidence.

In diagrams, the links of premises which 'hold' have solid lines and the links of premises which do not hold have dashed lines. Similarly, the link from an argument to its conclusion is displayed using a solid line if the argument is defensible, and shown with a dashed line if it is not defensible.

If both the pro argument, a1, and con argument, a2, had been acceptable in this example, the acceptability of the conclusion, "Nixon is a dove." would have depended, using the 'best argument' proof standard, on which argument is 'better', i.e. has priority. The Carneades model of argument does not tell us how to prioritize arguments. It merely assumes that users can do this in some way. The ordering may depend on the user or 'audience' [Bench-Capon and Sartor, 2003]. For example, one application scenario for a diagramming tool based on Carneades would be to help juries to reach a verdict, during their deliberations. In this scenario, the jury would deliberate about how to order the evidence, represented as arguments, and the software would visualize the consequences of various, proposed or chosen orderings.

Figure 2 presents a second example, a reconstruction of Toulmin's leading example [Toulmin, 1958], about whether or not Harry is a British citizen, illustrating some further conventions for diagramming Carneades argument graphs.

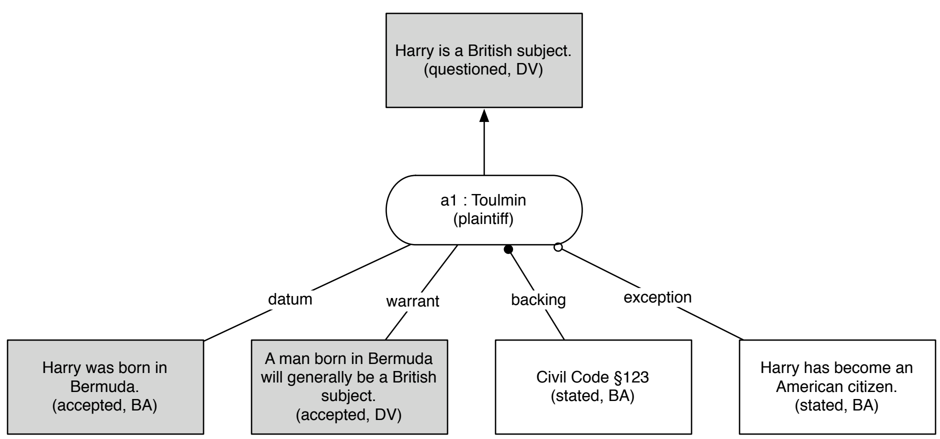

Figure 2: Reconstruction of Toulmin Diagrams in Carneades

The first thing to notice about this example is that it demonstrates how this diagramming method for Carneades generalizes Toulmin diagrams. Whereas arguments instantiating any argumentation scheme can be diagrammed with our method, Toulmin diagrams are specific to one particular argumentation scheme. Indeed, our method has in common with the Beardsley/Freeman ‘standard’ argument diagramming method [Beardsley, 1950, Freeman, 1991], as it is implemented by Araucaria [Rowe et al., 2003], the ability to handle arguments instantiating any number of different argumentation schemes in a single diagram. This example also illustrates how exceptions and assumptions are diagrammed, using open and closed circular arrowheads on the links of premises.

3 Reconstructing Wigmore Charts in Carneades

This diagramming method for Carneades is rather verbose compared to some alternatives. Here are some diagramming options which might be suitable for some purposes:

- Statements can be represented by keys or identifiers, rather than full sentences.
- Properties and labels which are irrelevant for whatever purpose the diagramming is being used for can be omitted.
- Argument nodes can be displayed as small figures by suppressing the identifier, argumentation scheme name and other properties of the argument except for its premise and conclusion links. Different shapes can be used to distinguish the argumentation schemes applied.

When all of these options are selected, the resulting diagrams are as compact as any other argument diagramming method with which are are familiar, in particular Wigmore charts. Interactive software applications for diagramming argument graphs could allow users to select among these options to control the compactness of the diagrams.

To provide a further example of Carneades diagrams, show how Carneades can be used to model reasoning with evidence in legal cases and compare our diagramming method with Wigmore charts, we now present a reconstruction of Wigmore’s chart of the evidence in the case of Commonwealth v. Umilian (1901, Supreme Judicial Court of Massachusetts, 177 Mass. 582), in the version presented by Bex, et al. [Bex et al., 2003], as shown in Figure 3.

In Wigmore charts, testimonial, circumstantial, explanatory and corroborative evidence are represented by squares, circles, triangles and angles, respectively. The number next to each figure is a key to a table with a list of statements. Each arrow or link represents an inference from one piece of evidence to another. A double bar at the top of a figure indicates the evidence was submitted by the defendant, rather than the plaintiff or prosecution. A dot in the center of a figure indicates that the evidence is accepted as a fact, by the user of the diagram. An infinity symbol below a figure shows the piece of evidence has been accepted as a fact by the court (judicially noticed). Some conventions in Wigmore diagrams seem redudant: explanatory evidence, submitted to ‘explain away’ other evidence, is both represented as an angle and, redundantly, placed to the left of the evidence to be explained away. Conversely, corroborative evidence, submitted to support the inference of other evidence, is represented as a triangle and, redundantly, placed to the right of the evidence to be supported. The figures for circumstantial and testimonial evidence are, redundantly, placed below the evidence they

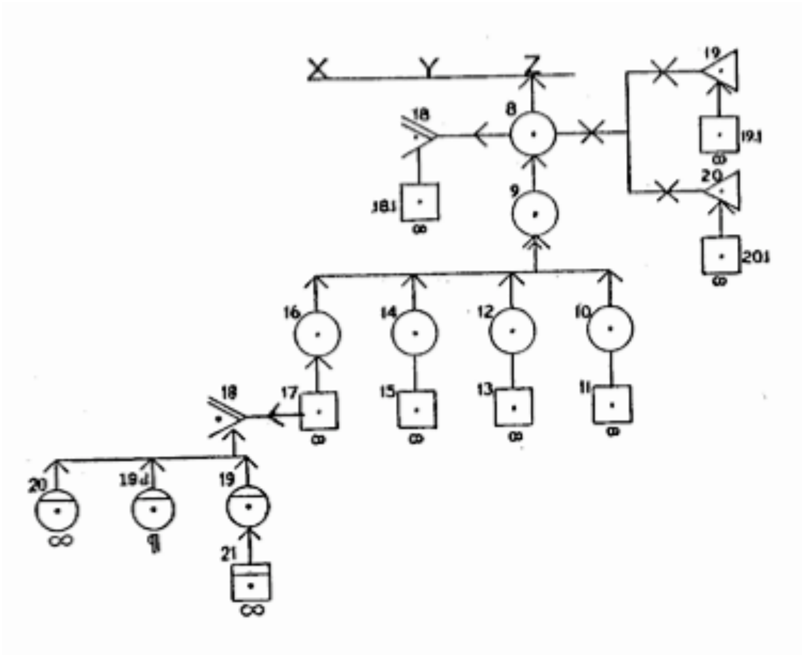

Figure 3: Wigmare Chart for the Umilian Case

tend to prove (affirmative evidence) or disprove (negatory evidence). The direction of the evidence, affirmative or negatory, is indicated by the bottom line of the figure for the evidence: negatory evidence is indicated by omitting the bottom line of the figure. (All evidence in the example diagram is affirmative.) Finally, Wigmore diagrams allow both the strength of belief in a piece of evidence and the strength of an inference from one piece of evidence to another to be indicated using labels on figures and arrows, respectively. Both of these strengths are selected from a small set of three or four ordinal values each. In the example chart, strong inferences are labeled with a X. The strengths of beliefs are not distinguished in this example. The capital letters at the top of the chart (X, Y, and Z) each represent a statement for some 'ultimate probandum' of the case. In the example, Z represents the charge that 'U killed J.'. Several pieces of evidence can be joined together in a single inference. This is indicated by a horizontal line joining the inferences. An example in Figure 3 is the inference from nodes 10, 12, 14 and 16 to node 9.

Figure 4 shows a reconstruction of this Wigmore chart in Carneades, making use of the options for producing more compact diagrams presented at the beginning of this section.

To make it easier to compare these two diagrams, we have laid out the nodes in the Carneades version in roughly the same way as in the Wigmore chart. The shapes chosen for argument nodes, indicating the argumentation scheme applied, are meant to be reminiscent of their counterparts in the Wigmore chart, with the exception of the use

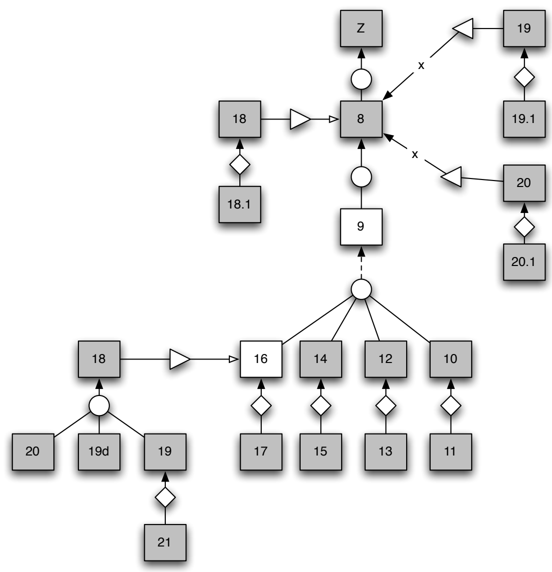

Figure 4: Carneades Diagram for the Umilian Case

of a diamond, instead of a square, for testimonial evidence, since rectangles are used for statements in our method.

The key idea in this reconstruction of Wigmore charts in Carneades is to represent the various types of evidence with different argumentation schemes. Testimonial evidence, for example, can be understood as an instantiation of a scheme for arguments from testimonial evidence.[4]

In the Wigmore diagram, all the pieces of evidence are 'believed', as indicated by the dots in each figure, and all the leaf nodes have been accepted by the court as judicial facts, as indicated by the infinity symbol below each leaf node. Carneades does not distinguish between statements which have been accepted by the user of the model and those accepted by the court. If we make the simplifying assumption that the court is the user of the diagram, these two types of acceptance could be represented uniformly in Carneades.

To determine the acceptability of statements using Carneades in this example, we first needed to assign proof standards. We have assumed the best argument standard (BA) applies uniformly to all statements and that arguments are ordered using a two-valued ordinal scale, as in the Wigmore chart. (The stronger arguments have been indicated by labeling the link to their conclusion with an 'x'.) To make things more interesting, we assume only the leaf nodes have been accepted, unlike in the original

[4]Wigmore's concept of circumstantial evidence is controversial. In [Walton, 2002], Walton argues that Wigmore's concept of circumstantial evidence covers diverse schemes for arguments from many kinds of evidence, other than witness testimony. Our diagramming method for Carneades is independent of particular argumentation schemes or their formulation. We interpret circumstantial evidence as an argumentation scheme in this article only to illustrate how Wigmore charts can be reconstructed using our system.

Wigmore chart, where all pieces of evidence have been accepted. As can be readily seen, only two statements are neither acceptable nor accepted, statements 9 and 16. Statement 16 is not acceptable, given the best argument standard, because it has a defensible con argument (with premise 18) which is just as strong as its strongest defensible pro argument (with premise 17). This in turn causes statement 9 to not be acceptable, since it is not supported by any defensible argument. Its only pro argument includes premise 16, which as we've just noted is not acceptable. Since we have assumed statement 16 to be at issue, rather than accepted as in the original Wigmore chart, premise 16 does not hold. Statement 8 is acceptable even though one of its pro arguments fails (the one with premise 9) because it is supported by two further pro arguments (with premises 19 and 20), both of which are defensible and each of which is stronger than the only defensible con argument (with premise 18). Thus the main claim, Z, is also acceptable, given the best argument standard, as it is supported by a defensible pro argument (with premise 8) and there are (currently) no arguments con Z, let alone defensible ones.

4 Conclusion

The diagramming method for Carneades presented here provides most of the features of Wigmore diagrams, but is founded on a formal, mathematical model of argument structure and evaluation which reflects the state of the art of the theory of argumentation in the field of philosophy. This diagramming method supports 'sense-making' application scenarios [Kirschner et al., 2002], which use argument visualization to help users to structure and understand arguments and their relationships, but goes beyond them by using the formal semantics of Carneades to inform users about the acceptability of statements.

Notice that Carneades is a qualitative model of argument, not a quantitative model such as probability theory. Thus, objections to probabilistic approaches to modeling legal evidence [Twining and McCrudden, 2006] may not directly apply to Carneades.[5] It remains for future work to compare Carneades with probability theory, in particular Bayesian Networks, as well as to investigate whether computational models of argument, and Carneades in particular, can overcome the objections which have been raised to mathematical models of legal evidence based on probability theory.

References

[Beardsley, 1950] Beardsley, M. C. (1950). *Practical Logic*. Prentice Hall, New York.

[Bench-Capon and Sartor, 2003] Bench-Capon, T. and Sartor, G. (2003). A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150(1–2):97–143.

[5] There are however qualitative versions of probability theory, [Parsons, 2001].

[Bex et al., 2003] Bex, F., Prakken, H., Reed, C., and Walton, D. (2003). Towards a formal account of reasoning with evidence: Argumentation schemes and generalizations. *Artificial Intelligence and Law*, 11(2-3):125–165.

[Freeman, 1991] Freeman, J. B. (1991). *Dialectics and the Macrostructure of Arguments: A Theory of Argument Structure*. Walter de Gruyter, Berlin / New York.

[Gordon, 2005] Gordon, T. F. (2005). A computational model of argument for legal reasoning support systems. In Dunne, P. E. and Bench-Capon, T., editors, *Argumentation in Artificial Intelligence and Law*, IAAIL Workshop Series, pages 53–64. Wolf Legal Publishers.

[Gordon et al., 2007] Gordon, T. F., Prakken, H., and Walton, D. (2007). The Carneades model of argument and burden of proof. *Artificial Intelligence*. In Press.

[Kirschner et al., 2002] Kirschner, P. A., Shum, S. J. B., and Carr, C., editors (2002). *Visualizing Argumentation: Software Tools for Collaborative Education and Sense-Making*. Springer-Verlag, London.

[Parsons, 2001] Parsons, S. (2001). *Qualitative Methods for Reasoning Under Uncertainty*. MIT Press.

[Rowe et al., 2003] Rowe, G. W. A., Reed, C. A., and Katzav, J. (2003). Araucaria: Marking up argument. In *European Conference on Computing and Philosophy*, Glasgow.

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

[Twining and McCrudden, 2006] Twining, W. and McCrudden, C., editors (2006). *Rethinking Evidence: Exploratory Essays*. Cambridge University Press, New York, 2nd edition.

[Walton, 2002] Walton, D. (2002). *Legal argumentation and evidence*. Pennsylvania State University Press, University Park, PA.

[Walton, 2006] Walton, D. (2006). *Fundamentals of Critical Argumentation*. Cambridge University Press.