Carneades User Manual

Tom Gordon

October 14, 2012

Contents

1 Getting Started

The Carneades Argumentation System is open source software, freely available for downloading at http://carneades.github.com.

Carneades provides web-based, collaborative software tools for:

- summarizing the arguments of a debate in an *argument graph*
- visualizing, browsing and navigating argument graphs
- critically evaluating arguments
- forming opinions, participating in polls and ranking stakeholders by the degree to which they share your views
- obtaining clear explanations, using argument graphs, of the different effects of alternative policies in particular cases

In this manual, we distinguish three kinds of users:

Public Users Users interested in understanding arguments and policies, whom we assume have had *no specialist training* in argumentation or policy modeling.

Analysts Users who have had *specialist training* in argumentation or rule-based policy modeling.

System Administrators Users with sufficient computer skills to install, configure and administer the Carneades software.

These are only roles. Some people may be able to use the system in more than one of these roles. In particular, we aim to make the software simple to install and administer, so that ultimately anyone with basic computer skills will be able to serve as "system administrator".

The rest of this manual is organized in the following chapters:

- Argument Graphs
- Browsing, Visualizing and Evaluating Arguments
- Forming, Polling and Comparing Opinions
- Analyzing and Comparing Policies
- Editing Argument Graphs

- Modeling Policies and Argumentation Schemes
- System Administration
- Application Scenarios

The chapter on argument graphs provides a concise and easy overview of the underlying data model used by all the tools and is recommended to be read first by all users. The next three chapters are for public users and explain how to use the tools for argument browsing, opinion polls and policy analysis, respectively. The chapters on editing argument graphs and modeling policies are for analysts with some prior knowledge of argument reconstruction methods and rule-based systems. The system administration chapter is recommended reading for anyone wishing to install, configure and administer the Carneades software. Finally, the chapter on application scenarios may be of interest to managers and public administrators looking for ways to improve the quality of decision-making processes in their organizations.

2 Argument Graphs

Argument graphs model relationships among arguments, statements and source documents. To build the model, source documents need to be interpreted to find the arguments, in a process called “argument reconstruction”.

Consider the following simple legal argument: Johnny violated the law by riding his skateboard in the park.

This same argument can be expressed in many different ways. Here are some examples:

1. Because Johnny rode his skateboard in the park he violated the law.
2. Vehicles are prohibited in the park. Someone who does something which is prohibited violates the law. Johnny rode his skateboard in the park. A skateboard is a vehicle. Therefore Johnny violated the law.
3. Johnny hat gegen das Gesetz verstoßen, weil er sein Skateboard im Park benutzt hat.

The first example just reordered the premise and the conclusion of the argument, putting the premise first. The second example reveals some implicit premises of the original formulation of the argument. The third example is a German translation of the original formulation of the argument.

All four of these texts, including the original formulation, express the *same* argument, but in different ways. In a large-scale debate, for example about

European policy issues, the same argument might be expressed in *thousands* of different ways in many different languages. (The European Union has 23 official languages.)

One of the purposes of argument graphs is to provide a way to summarize complex debates with exactly one node in the graph for each argument put forward in the debate. A single argument graph is used to represent all the arguments put forward in a debate, from all participants. The nodes can quote one or more source documents, and include links to these source documents, so no information need be lost and all voices can still be heard, inclusively. Grouping the different formulations of an argument together into a single node in the graph, abstracting away details, makes it possible to quickly obtain an overview of the arguments and to obtain a clearer picture of relationships among arguments. A hypertext or map of the source documents directly, without an argument graph, would make it difficult to "see the forest for the trees".

2.1 Data Model

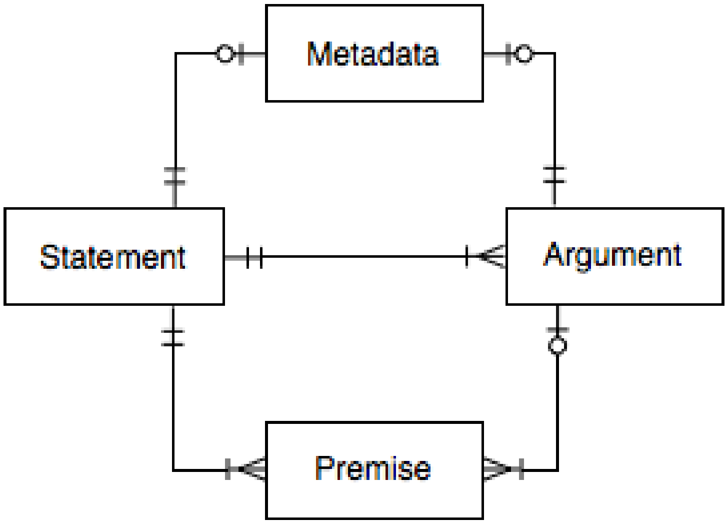

Figure 1: Entity-Relationship Diagram

The entity-relationship diagram above shows the elements of argument graphs and their connections. (The figure does not visualize a particular argument graph, but rather relationships between the elements of argument graphs in general.)

The two main elements of argument graphs are statements and arguments. Statements represent propositions, claims and assertions. Arguments represent simple inferences from one or more premises to a single conclusion. Again, there should be only one statement or argument in the graph for each statement and argument in the source documents, no matter how many different ways the statement or argument has been expressed in source documents. Some or all formulations of the statement or argument can be quoted or referenced in the metadata of the statement or argument node. See the discussion of metadata below for further information.

As can be seen in the entity-relationship diagram, arguments are linked to statements in two ways in argument graphs. Each argument has exactly one conclusion, which is a statement, and zero or more premises, where each premise has exactly one statement node. A statement may the conclusion or premise of more than one argument.

A statement may be both a conclusion and a premise, resulting in complex argument graphs, representing chains or trees of reasoning. Argument graphs may contain cycles. A simple cycle would result if a statement and a premise of the same argument. There are methods for resolving these cycles when evaluating argument graphs.

A statement in an argument graph represents both a proposition and the negation of the proposition. To continue with our example, the sentences "Johnny rode his skateboard in the park" and "Johnny did not ride his skateboard in the park" would be represented by a single statement in an argument graph. Conclusions and premises of argument can negated using con arguments and negative premises, respectively. That is, there are two kinds of arguments, pro and con. An argument is pro if its conclusion claims the statement is true and con if it claims the statement is false. Similarly, there are two kinds of premises, positive and negative. A positive premise holds if its statement is presumably true ("in"). Conversely, a negative premise holds only if its statement is presumably false ("out").

Prior models of argument graphs do not distinguish pro and con arguments or positive and negative premises. Rather, in these prior approaches all argument nodes are pro and all premises are positive. Our approach has the advantage of reducing the number of statements up to 50%, resulting in more compact summaries of the arguments.

2.2 Statement Properties

id A Uniform Resource Name (URN) serving as a unique identifier for the statement, world-wide.

weight A real number in range of 0.0-1.0 representing the degree to which the statement is accepted as true by the users, where 0.0 means the statement is *rejected* (believed to be false by the users) and 1.0 means the statement is *accepted* (believed to be true). This information is collected via polls.

proof standard The method used to combine pro and con arguments. Several proof standards are supported by the system. For most purposes, the "preponderance of the evidence" standard should suffice.

value A real number in the range 0.0-1.0, storing the output of the argument evaluation process, where 0.0 means the statement is *out* (presumably false), 1.0 means the statement is *in* (presumably true) and all other values mean the arguments are insufficient for making any presumptions about the truth or falsity of the statement.

atom An optional formal representation of the statement in predicate logic. (This feature is for analysts and need not interest public users.)

main A Boolean value (true or false) used to indicate whether the statement is one of the main issues of the debate modeled by the argument graph.

2.3 Argument Properties

id A Uniform Resource Name (URN) serving as a unique identifier for the argument, world-wide.

direction Pro or con.

strict A Boolean value (true or false) expressing whether the conclusion of the argument is necessarily true when its premises are true (strict arguments) or only presumably true. Nonstrict arguments are called "defeasible" arguments.

scheme The name of the argumentation scheme applied, if any. Optional. Examples: "argument from credible source", "argument from practical reasoning".

weight A real number in range of 0.0-1.0, representing the relative weight of the argument, compared to other arguments pro and con the conclusion of the argument. This information is collected via polls.

value A real number in in the range 0.0-1.0, used to record the output of the argument evaluation process, where 0.0 means the argument is *out* (not acceptable), 1.0 means the statement is *in* (acceptable) and all other values mean the arguments in the graph, taken together, are insufficient for determining the acceptability of this argument.

2.4 Premise Properties

polarity Positive or negative.

role The role of the premise in the argumentation scheme applied. Examples: "minor", "major".

implicit A Boolean value (true or false). Can be used to note that the premise was not explicit in the source documents from which the argument node was reconstructed.

2.5 Metadata

The argument graph as as whole, as well as each of its statements and arguments, can be annotated with metadata, using the Dublin Core. There are 15 elements in the Dublin Core. Each element may zero or more values. Here is a list of the Dublin Core elements:

1. Title
2. Creator
3. Subject
4. Description
5. Publisher
6. Contributor
7. Date
8. Type
9. Format
10. Identifier
11. Source
12. Language
13. Relation
14. Coverage
15. Rights

See Dublin Core for a detailed description and usage guidelines for each element. The Dublin Core is intended to be useful for describing a wide range of "resources" on the World-Wide Web. Not all of the elements may be applicable for argument graphs.

In addition, Carneades allows each metadata record to be assigned an optional "key", a string which can be used as a label to refer to the metadata record, such as “BenchCapon:2008”, similar to the way citation keys are used in bibliographic databases such as BibTeX.[1] At most one key should be provided.

Carneades provides special support for providing description elements of the Dublin Core in multiple languages (English, German, French, . . .) and for formatting these descriptions using the Markdown language. This feature can be used to include quotations from and links to source documents in the descriptions of both statements and arguments.

3 Browsing, Visualizing and Evaluating Arguments

4 Formulating, Polling and Comparing Opinions

5 Analysing and Comparing Policies

6 Editing Argument Graphs

7 Modeling Policies and Argumentation Schemes

8 System Administration

9 Applications Scenarios

[1] http://en.wikipedia.org/wiki/BibTeX