Goal Selection in Argumentation Processes

A Formal Model of Abduction in Argument Evaluation Structures

Stefan BALLNAT [1] and Thomas F. GORDON [2]
Fraunhofer FOKUS, Berlin

Abstract. When argumentation is conceived as a kind of *process*, typically a dialogue, for reasoning rationally with limited resources under conditions of incomplete and inconsistent information, arguers need heuristics for controlling the search for arguments to put foward, so as to move from stage to stage in the process in an efficient, goal-directed way. For this purpose, we have developed a formal model of abduction in argument evalution structures. An argument evaluation structure consists of the arguments of a stage, assumptions about audience and an assignment of proof standards to issues. A derivability relation is defined over argument evaluation structures for the literals 'in' a stage. Literals which are not derivable in a stage are 'out'. Abduction is defined as a relation between an argument evaluation structure and sets of literals, called 'positions', which, when the assumptions are revised to include the literals of the position, would make a goal literal in or out, depending of the standpoint of the agent. Soundness, minimiality, consistency and completeness properties of the abduction relation are proven. A heuristic cost function estimating how difficult it is to find or construct arguments pro a literal in the domain can be used to order positions and literals within positions. We compare our work to abduction in propositional logic, in particular the Assumption-Based Truth Maintenance System (ATMS).

Keywords. Abduction, Argumentation, Argument Evaluation, Dialogues, Heuristics, Process Models, Proof Standards, Relevance

1. Introduction

We view argumentation as a kind of process for reasoning rationally about problems which are not well-formed or semi-decidable with incomplete or inconsistent information and limited computational resources [1,2]. Persuasion dialogues [3] between a proponent and respondent about some claim or thesis are the prototypical type of argumentation process, but argumentation processes in our conception are not restricted to dialogues, in their usual sense as conversations between two or more persons.[3]

[1] stefan.ballnat@fokus.fraunhofer.de
[2] thomas.gordon@fokus.fraunhofer.de

[3] Dialogues in philosophy and AI are often generalized to cover as well the reasoning processes of single agents, switching between pro and con roles. Presumably most if not all argumentation processes can be viewed as dialogues in this generalized sense.

Figure 1 shows a simple argumentation process, where an *agent* is preparing his case, to be presented later to the audience, by constructing arguments from information found using some *information service*.

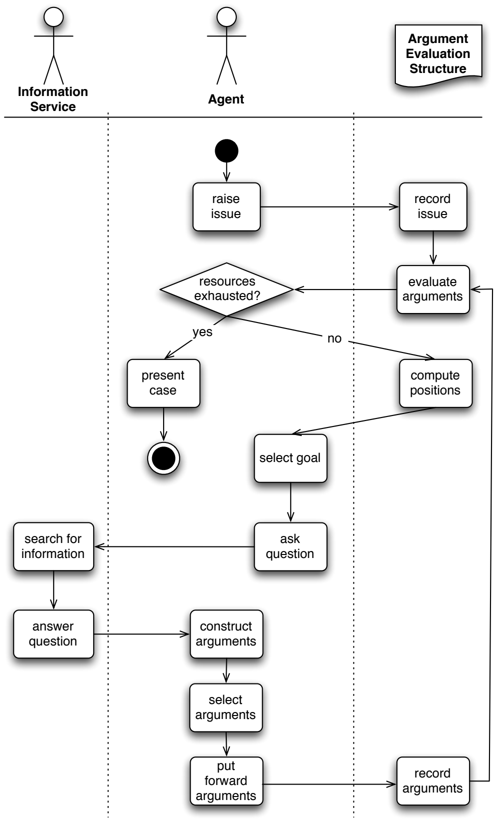

Figure 1. A Simple Argumentation Process

Arguments are put forward, recorded and evaluated in an *argument evaluation structure*, taking into consideration applicable *proof standards* [4,5,6] and assumptions about the *audience*. The agent needs to formulate some kind of impression of his audience, in order to avoid expending resources trying to prove propositions which the audience would accept without proof and also to select arguments which are likely to be persuasive.

Each time through the loop, if resources remain, the agent will use the argument evaluation structure to abduce alternative *positions*, where, similar to [7], a position is a set of propositions which, if added to the assumptions, would make the main issue provable (in) or not provable (out), depending of the standpoint of the agent. From these alternative positions, a subgoal is then selected to work on next, taking into consideration the estimated cost of proving each proposition of a position.

Our focus here is on the task of selecting the next goal to work on. This is an essential, central task of all argumentation processes. An agent cannot argue well without some means of efficiently choosing subgoals which are likely to lead to arguments which are effective for helping to resolve the main issue in a favorable way, given the agent's standpoint.

For the purpose of supporting goal selection, this paper presents a formal model of abduction in argument evaluation structures. Abduction has several meanings in logic. It can mean a method for inferring explanations of observations, but we are using the term by analogy to a more formal meaning, where abduction is one of three kinds of inference relations, together with deduction and induction.

The rest of the paper is organized as follows. The next section presents the formal model. This is followed with a section presenting an example illustrating the model. The paper closes with a discussion of related and possible future work.

2. The Formal Model

The model of goal selection builds on our prior work on argument evaluation structures [5,6]. To make this paper self-contained, we begin by summarizing this prior work.

We begin with the concept of an argument. Informally, an argument is a structure linking a set of premises to a conclusion. Some of the critical questions of the argumentation scheme used to construct the argument are modeled as exceptions of the argument.

Definition 1 (argument) *Let $\mathcal{L}$ be a propositional language. An **argument** is a tuple $\langle P, E, c \rangle$ where $P \subset \mathcal{L}$ are its **premises**, $E \subset \mathcal{L}$ are its **exceptions** and $c \in \mathcal{L}$ is its **conclusion**. For simplicity, c and all members of P and E must be literals, i.e. either an atomic proposition or a negated atomic proposition. Let p be a literal. If p is c, then the argument is an argument **pro** p. If p is the complement of c, then the argument is an argument **con** p.*

Arguments do not need to be deductively valid. For example, premises needed to make the argument deductively valid can be left implicit. (Such arguments are called 'enthymemes'.) An argument is dialectically valid only if it furthers the goals of the argumentation process [3]. One way to assess the dialectically validity of an argument is to check whether it is an instance of an argumentation scheme which is accepted by the procedural rules (protocol) of the particular argumentation process in the problem domain.

Argumentation is viewed as a process. To fully model different kinds of proof burdens, it is useful to divide the process into three phases, an open, argumentation and closing phase and to distinguish between claimed and questioned propositions. But for our purposes here these details are not necessary.

Definition 2 (argumentation process) *An **argumentation process** is a sequence of **stages** where each stage is a set of arguments. In every chain of arguments, $a_1, \ldots a_n$, constructable from the arguments in a stage by linking the conclusion of an argument to a premise or exception of another argument, a conclusion of an argument a_i may not be a premise or exception of an argument a_j, if $j < i$. A set of arguments which violates this condition is said to contain a* cycle *and a set of arguments which complies with this condition is called* cycle-free*.*

Notice that arguments both pro and con some proposition can be included in the arguments of a stage without causing a cycle.

Next we need a structure for evaluating arguments, to assess the acceptability of propositions at issue. As in value-based argumentation frameworks [8,9] arguments are evaluated with respect to assumptions about an *audience*.

Definition 3 (audience) *An* audience *is a structure* $\langle \Phi, f \rangle$*, where* $\Phi \subset \mathcal{L}$ *is a consistent set of literals assumed to be acceptable by the audience and f is a partial function mapping arguments to real numbers in the range* $0.0 \ldots 1.0$*, representing the relative weights assumed to be assigned by the audience to the arguments.*

An argument evaluation structure associates an audience with a stage of process and assigns proof standards to issues, providing a basis for evaluating the acceptability of propositions to this audience.

Definition 4 (argument evaluation structure) *An* ***argument evaluation structure*** *is a tuple* $\langle \Gamma, \mathcal{A}, g \rangle$*, where* Γ *is a stage in a argumentation process,* $\mathcal{A}$ *is an audience and* g *is a total function mapping propositions in* $\mathcal{L}$ *to their applicable proof standards in the process. A* ***proof standard*** *is a function mapping tuples of the form* $\langle p, \Gamma, \mathcal{A} \rangle$ *to the Boolean values true and false, where* p *is a literal in* $\mathcal{L}$*.*

Given an argument evaluation structure, the acceptability of a proposition is defined by its proof standard.

Definition 5 (acceptability) *Let* $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ *be an argument evaluation structure, where* $\mathcal{A} = \langle \Phi, f \rangle$*. A literal* p *is* ***acceptable*** *in* $\mathcal{S}$ *if and only if* $g(p)(p, \Gamma, \mathcal{A})$ *is true.*

Derivability in an argument evaluation structure can then be defined as a kind of nonmonotonic inference relation as follows:

Definition 6 (derivability) *Let* $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ *be an argument evaluation structure, where* $\mathcal{A} = \langle \Phi, f \rangle$*. A literal* p *is* ***in*** S*, denoted* $(\Gamma, \Phi) \vdash_{f,g} p$*, if and only if*

- $p \in \Phi$ *or*
- $(\neg p \notin \Phi$ *and* p *is acceptable in* $S)$

Otherwise p *is* ***out****, denoted* $(\Gamma, \Phi) \nvdash_{f,g} p$*.*

Obviously much of the work of argument evaluation has been delegated to the proof standards. All the proof standards we have defined make use of the concept of argument applicability, so let us define this concept first.

Definition 7 (argument applicability) *Let* $S = \langle \Gamma, \mathcal{A}, g \rangle$ *be an argument evaluation structure. An argument* $\langle P, E, c \rangle$ *is* ***applicable*** *in this argument evaluation structure if and only if*

- *the argument is a member of* Γ*, the arguments of the stage*
- *every proposition* $p \in P$ *is in*
- *every proposition* $p \in E$ *is out*

In [6] we defined five proof standards, most of them modeling legal proof standards: 1) scintilla of the evidence, 2) preponderance of the evidence, 3) clear and convincing evidence, 4) beyond a reasonable doubt and 5) dialectical validty. To illustrate proof standards we present here the definition of two of these standards, dialectical validity and preponderance.

Definition 8 (dialectical validity) *Let $\langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and let p be a literal in $\mathcal{L}$. dv$(p, \Gamma, \mathcal{A})$ = true if and only if there is at least one applicable argument pro p in Γ and no applicable argument con p in Γ.*

Definition 9 (preponderance of the evidence) *Let $\langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and let p be a literal in $\mathcal{L}$. pe$(p, \Gamma, \mathcal{A})$ = true if and only if*

- *there is at least one applicable argument pro p in Γ and*
- *the maximum weight assigned by the audience $\mathcal{A}$ to the applicable arguments pro p is greater than the maximum weight of the applicable arguments con p.*

Given these preliminaries we can now turn to our primary task, of defining abduction and computing goals. We begin by the defining the *labels* of literals and arguments in an argument evaluation structure, where each label is a propositional formula. Each literal has two labels. The labels express conditions which, if accepted by the audience, would make the literal in or, respectively, out. Similarly, arguments also have two labels, expressing conditions which, if accepted by the audience, would make the argument applicable or, respectively, not applicable.

The base case in the definition of the labels of literals covers the case where the literal is assumed to have been accepted or rejected by the audience. If however the literal has not been assumed to have been accepted or rejected by the audience, then the label is a function of the proof standard assigned to the literal.

Definition 10 (literal labels) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ an argument evaluation structure with $\mathcal{A} = \langle \Phi, f \rangle$. The* in *and* out *labels of a literal p are defined as:*

- $\textit{in-label}(p, \mathcal{S}) := \begin{cases} \top & \textit{if } p \in \Phi \\ p & \textit{if } \neg p \in \Phi \\ \textit{ps-in-label}(p, \mathcal{S}) & \textit{otherwise} \end{cases}$
- $\textit{out-label}(p, \mathcal{S}) := \begin{cases} \neg p & \textit{if } p \in \Phi \\ \top & \textit{if } \neg p \in \Phi \\ \textit{ps-out-label}(p, \mathcal{S}) & \textit{otherwise} \end{cases}$

where the helping functions ps-in-label and ps-out-label are defined as follows:

- $\textit{ps-in-label}(p, \mathcal{S}) := \begin{cases} \textit{dv-in-label}(p, \mathcal{S}) & \textit{if } g(p) = \textit{dv} \\ \textit{ba-in-label}(p, \mathcal{S}) & \textit{if } g(p) = \textit{pe} \end{cases}$
- $\textit{ps-out-label}(p, \mathcal{S}) := \begin{cases} \textit{dv-out-label}(p, \mathcal{S}) & \textit{if } g(p) = \textit{dv} \\ \textit{ba-out-label}(p, \mathcal{S}) & \textit{if } g(p) = \textit{pe} \end{cases}$

To save space, the definitions of ps-in-label and ps-out-label handle only the two proof standards defined above, preponderance of evidence and dialectical validity, but they can be extended in a straightforward manner to handle other proof standards.

Before defining the labeling functions for these two specific proof standards we first define argument labels. Using the definition of argument applicability, the in-label of arguments is the conjunction of all in-labels of its premises and of all the out-labels of its exceptions. The out-label of an argument is simply the negation of its in-label, where the negation of a literal's in-label is the literal's out-label and vice versa, not the negation of the literal represented by the in-label.

Definition 11 (argument labels) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and $a = \langle P, E, c \rangle \in \Gamma$ be an argument. We define the* ***argument labels*** *for a as follows:*

- $\textit{in-label}(a, \mathcal{S}) := \bigwedge_{p \in P} \textit{in-label}(p, \mathcal{S}) \wedge \bigwedge_{e \in E} \textit{out-label}(e, \mathcal{S})$
- $\textit{out-label}(a, \mathcal{S}) := \bigvee_{p \in P} \textit{out-label}(p, \mathcal{S}) \vee \bigvee_{e \in E} \textit{in-label}(e, \mathcal{S})$

Next, we present the labeling functions for the two proof standards, using the argument labels, beginning with dialectical validity. The in-label for a literal assigned the dialectical validity standard is defined as the conjunction of the out-labels of all its con-arguments and the disjunction of the in-labels of all its pro-arguments. We also add the literal itself as a disjunct to the label in order to later enable goals to be derived for all literal nodes in an argument graph and not just leaves. Notice that the dialectical validity out-label is almost the negation of its in-label.

Definition 12 (dialectical validity label) *Let $S = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and $p \in \mathcal{L}$ a literal. The* ***dialectical validity label*** *for p in S is defined as follows:*

- $\textit{dv-in-label}(p, \mathcal{S}) := (\bigvee_{a \in Pro} \textit{in-label}(a, \mathcal{S}) \wedge \bigwedge_{a \in Con} \textit{out-label}(a, \mathcal{S})) \vee p$
- $\textit{dv-out-label}(p, \mathcal{S}) := \bigwedge_{a \in Pro} \textit{out-label}(a, \mathcal{S}) \vee \bigvee_{a \in Con} \textit{in-label}(a, \mathcal{S}) \vee \neg p$

where Pro and Con denote the sets of pro- and con-arguments for p in Γ.

Coming to the second proof standard, preponderance of the evidence, we must consider the weights of the arguments. The label of a literal assigned this standard is the disjunction of the in-labels of all its pro-arguments (assuring the existence of an applicable pro-argument) combined with the conjunction of the out-labels of all its con-arguments with greater or equal weight (assuring all applicable con-arguments have less weight). We again add the literal itself to the label. For the out-label we again turn the tables and require that either every pro-argument be not applicable or that there exist an applicable con-argument of greater or equal weight.

Definition 13 (preponderance label) *Let $S = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and $p \in \mathcal{L}$ a literal. The* ***preponderance label*** *for p in S is defined as follows:*

- $\textit{ba-in-label}(p, \mathcal{S}) := (\bigvee_{a \in Pro} [\textit{in-label}(a, \mathcal{S}) \wedge \bigwedge_{\substack{b \in Con, \\ f(a) \leq f(b)}} \textit{out-label}(b, \mathcal{S})]) \vee p$

- *ba-out-label*$(p, \mathcal{S}) := (\bigwedge_{a \in Pro} [$*out-label*$(a, \mathcal{S}) \vee \bigvee_{\substack{b \in Con, \\ f(a) \leq f(b)}}$ *in-label*$(b, \mathcal{S})]) \vee \neg p$

where Pro and Con denote the sets of pro and con arguments for p *in* Γ.

It is easy to see that the out-label of a literal is in general not the same as the in-label of the literal's complement. To make this clear, consider a small example of an argument evaluation structure $\mathcal{S}$ with just one literal q, no arguments and no assumptions for the audience. No matter what proof standard is assigned to q we obtain out-label$(q, \mathcal{S}) = \top \vee \neg q$ for the out-label whereas the in-label for the complement of q is in-label$(\neg q, \mathcal{S}) = \neg q$.

Labels for further proof standards can be defined similarly.

A label of a statement is transformed into a set of positions by first reducing the label to minimal disjunctive normal form and then interpreting the clauses of the formula as alternative positions consisting of the corresponding literals.

Definition 14 (position sets) *Let* l *be a label of a literal* p *in an argument evaluation structure* $\mathcal{S}$. *Let* $\lambda_l = C_1 \vee \ldots \vee C_n$ *be a formula equivalent to* l *in minimal disjunctive normal form. We define the* ***position set*** *of the label, denoted ps*(l)*, as follows: ps*$(l) = \{gl_{C_i} \mid 1 \leq i \leq n\}$ *where* $gl_{C_i} = \{L_{ij} \mid 1 \leq j \leq m\}$ *for* $C_i = L_{i1} \wedge \ldots \wedge L_{im}$

Abduction is defined as a relation between an argument evaluation structure and a position, which holds if adding the literals of the position to the assumptions about the audience would make a goal literal in or out, depending of the standpoint of the agent.

Definition 15 (abduction) *Let* $\mathcal{S}$ *be an argument evaluation structure and let* $p \in \mathcal{L}$ *be a literal. We define two abductive inference relations,* $\Vdash_{in}$ *and* $\Vdash_{out}$ *as follows.*

- $(p, \mathcal{S}) \Vdash_{in} \Delta$ *if and only if* $\Delta \in ps($*in-label*$(p, \mathcal{S}))$
- $(p, \mathcal{S}) \Vdash_{out} \Delta$ *if and only if* $\Delta \in ps($*out-label*$(p, \mathcal{S}))$

We have proved that these abduction relations are sound, minimal, consistent and complete with respect to the underlying derivability inference relation. There is space only to present sketches of the proofs of these properties here. The full proofs are available in a technical report [10].

First we need to take care of a technical issue, namely aggregating two positions in such a way that the literals of one position replace the complementary literals of the other position. We need this operator to revise the assumptions the agent makes about literals accepted by the audience after persuading the audience to accept a position.

Definition 16 (assumption revision) *Let* Φ *be the set of literals assumed to be accepted by the audience and let* Δ *be a position which the agent wants to persuade the audience to accept. The* ***revised assumptions*** *about the audience are* $\Phi \uplus \Delta = \{L \mid L \in \Delta \vee (L \in \Phi \wedge \neg L \notin \Delta)\}$

By soundness, we mean that revising the assumptions about the audience to include the literals of a position makes the issue in or out in the resulting argument evaluation structure, depending on whether the agent is interested in proving or disproving the literal at issue. This is the most important property, because it assures the agent that persuading the audience to accept the literals of a position should be effective in persuading the audience to accept the agent's standpoint with respect to the literal at issue.

Theorem 1 (soundness) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure with an audience $\mathcal{A} = \langle \Phi, f \rangle$. Let $p \in \mathcal{L}$ be a literal and Δ be a position. The following statements hold:*

- $(p, \mathcal{S}) \Vdash_{in} \Delta \Rightarrow (\Gamma, \Phi \uplus \Delta) \vdash_{f,g} p$
- $(p, \mathcal{S}) \Vdash_{out} \Delta \Rightarrow (\Gamma, \Phi \uplus \Delta) \nvdash_{f,g} p$

The proof is straight-forward because we defined the statement and argument labels on the basis of the acceptability and applicability definitions. Assumption revision preserves the consistency of the assumptions of the audience.

Theorem 2 (minimality) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure and $p \in \mathcal{L}$ be a literal. For all Δ_1 and Δ_2, if $(p, \mathcal{S}) \Vdash_{in} \Delta_1$ and $(p, \mathcal{S}) \Vdash_{in} \Delta_2$ or if $(p, \mathcal{S}) \Vdash_{out} \Delta_1$ and $(p, \mathcal{S}) \Vdash_{out} \Delta_2$, then $\Delta_1 \not\subset \Delta_2$.*

Minimality follows from the use of minimal disjunctive normal form to transform labels into position sets.

Theorem 3 (consistency) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure, $p \in \mathcal{L}$ a literal and Δ a position. If $(p, \mathcal{S}) \Vdash_{in} \Delta$ or $(p, \mathcal{S}) \Vdash_{out} \Delta$ then Δ is a consistent set of literals in classical propositional logic.*

Consistency follows from the use of minimal disjunctive normal form to transform the label of an issue into a position set, since the literal of the issue is added as a disjunct of the formula.

Finally, completeness guarantees that all possible positions enabling the agent to prove or disprove the literal at issue are abduced.

Theorem 4 (completeness) *Let $\mathcal{S} = \langle \Gamma, \mathcal{A}, g \rangle$ be an argument evaluation structure with an audience $\mathcal{A} = \langle \Phi, f \rangle$. Let $p \in \mathcal{L}$ be a literal and Δ a position. The following statements hold:*

- $(\Gamma, \Phi \uplus \Delta) \vdash_{f,g} p \Rightarrow \exists \Delta' \subseteq \Delta . (p, \mathcal{S}) \Vdash_{in} \Delta'$
- $(\Gamma, \Phi \uplus \Delta) \nvdash_{f,g} p \Rightarrow \exists \Delta' \subseteq \Delta . (p, \mathcal{S}) \Vdash_{out} \Delta'$

The proof of completeness concentrates on the definition of literal labels, in particular its use of specific labels for each proof standard, to show that all sets of assumptions that make a literal acceptable (or not acceptable) are covered by its labels.

Having defined positions, we can heuristically select a subset of the positions, as part of a reasoning strategy, using a total order on positions. Such an order could be defined, e.g., using an estimate of the cost of proving each literal. This would enable the agent to select one of the cheapest positions to work on next.

Definition 17 (preferred positions) *Let $\mathcal{S}$ be an argument evaluation structure, $p \in \mathcal{L}$ a literal at issue, and $\leq$ a total order on positions. A position Δ for proving p, where $(p, \mathcal{S}) \Vdash_{in} \Delta$, is a **preferred position** if and only if $\forall \Delta' . (p, \mathcal{S}) \Vdash_{in} \Delta' \Rightarrow \Delta \leq \Delta'$.*

Similarly, if Δ is a position for disproving p, where $(p, \mathcal{S}) \Vdash_{out} \Delta$, then Δ is a preferred position if and only if $\forall \Delta' . (p, \mathcal{S}) \Vdash_{out} \Delta' \Rightarrow \Delta \leq \Delta'$.

Finally, we define two relevance properties of arguments. Intuitively, an argument is relevant for a proving a goal literal, if its conclusion is a member of a position for this literal. The stronger relevance property requires in addition that the conclusion of the argument be a member of a preferred position.

Definition 18 (relevance) *Let $a = \langle P, E, c\rangle$ an argument and $\mathcal{S}$ be an argument evaluation structure and $p \in \mathcal{L}$ be an issue.*

- *The argument a is **weakly relevant** in S if and only if $\exists\Delta.[(p, \mathcal{S}) \Vdash_{in} \Delta \vee (p, \mathcal{S}) \Vdash_{out} \Delta] \wedge c \in \Delta$*
- *The argument a is **strongly relevant** in S if and only if c is a member of a preferred position for p in $\mathcal{S}$.*

This notion of relevance is analogous to Sperber and Wilson's [11], who define relevance by two conditions. Weak relevance is similar to the first condition, requiring "large effects" on the "context". The context is here the argument evaluation structure; the large effect of proving a weakly relevant statement is to bring the agent "one step closer" to proving or disproving the main issue. Strongly relevance is analogous to the second condition of "small effort" because we take the total order on the positions into account to select the "cheapest" position.

3. Example

Figure 2 shows a small argument graph for which it is already hard to see what goal to choose next. In the argument evaluation structure $\mathcal{S} = \langle\Gamma, \mathcal{A}, g\rangle$ we have the stage $\Gamma = \{a_1, a_2, a_3, a_4, a_5\}$ and the audience $\mathcal{A} = \langle\Phi, f\rangle$, where $\Phi = \{\neg q, s, \neg t, \neg v, w\}$ and the weights assigned by f are as shown in the figure. The proof standard assigned by g is preponderance of the evidence for all literals. The issue is p.

In the figure, literals are shown in boxes and arguments in circles. The weights assigned by the audience to arguments are shown above the circles. Pro arguments are displayed with filled arrowheads; con arguments with open arrowheads. Premises are visualized by a solid line linking a proposition to an argument; Exceptions are visualized with dashed lines. By convention, only positive literals are displayed in boxes. Negative conclusions are visualized using con arguments. Negated premises are shown with a cross mark on the link between the premise and the argument, but there is no example in this figure. Literals assumed to have been accepted by the audience are shown with a check mark in the lower left corner of the box for the literal; literals assumed to have been rejected by the audience are shown with an X mark. (Acceptance of a literal P implies rejection of $\neg P$ and vice versa.)

Computing the labels for p and its complement we obtain:

- in-label$(p, \mathcal{S}) = (\top \wedge q \wedge (\top \vee v \vee t \vee \neg r)) \vee p \equiv q \vee p$
- out-label$(p, \mathcal{S}) = \neg s \vee \top \vee (u \wedge \top \wedge \top) \vee r \vee \neg p \equiv \top$
- in-label$(\neg p, \mathcal{S}) = (u \wedge \top \wedge \top) \vee r \vee \neg p \equiv u \vee r \vee \neg p$
- out-label$(\neg p, \mathcal{S}) = \top \vee v \vee t \vee \neg r \vee p \equiv \top$

The four corresponding sets of positions are:

- $\{q\}$ and $\{p\}$ if the agent wants to prove p

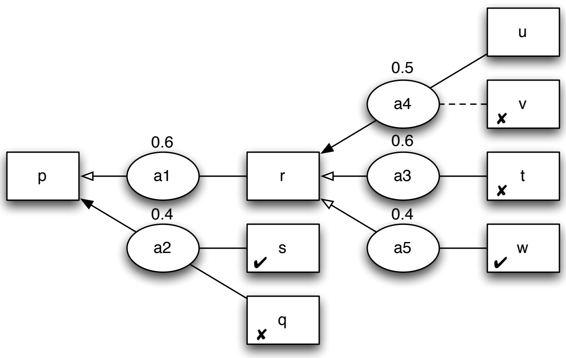

Figure 2. An example argument graph

- $\{\}$ if the agent wants to disprove p.
- $\{u\}$, $\{r\}$ and $\{\neg p\}$ if the agent wants to prove $\neg p$.
- $\{\}$ if the agent wants to disprove $\neg p$.

So, if the agent is interested in disproving p or $\neg p$ no further arguments should be needed to persuade the audience, since neither literal currently satisfies its proof standard. If however the agent wants to prove p, he can either try to construct another argument pro p or arguments sufficient to make q acceptable to the audience. The agent need not be concerned yet about rebutting the argument con p because it is not currently applicable.

4. Discussion

We have presented a model of abduction in an argument evaluation structure which enables a principled method for selecting goals to work on in argumentation processes, to help agents to construct arguments in an efficient, goal-directed way.

Again, we are using the term abduction by analogy to its formal meaning, as one of three kinds of inference relations, together with deduction and induction, not as a method for inferring causes of observations. The role of deduction is played by the derivability relation in our system, which infers literals which 'in' an argument evaluation structure. Induction in our system would be the inference of further arguments which would make a goal statement in if the arguments are added to the stage of the argument evaluation structure. Finally, abduction in this framework is the inference of literals which, if added to the set of literals assumed to be accepted by the audience, would make a goal literal derivable (in) in the argument evaluation structure.

This work was inspired by de Kleer's Assumption-Based Truth Maintenance System (ATMS) [12,13] and its application to controlling problem-solvers [14]. Whereas the ATMS performs abduction in a subset of classical propositional logic, our system performs abduction in an argumentation evaluation structure, which is a kind of nonmonotonic inference relation. Junker showed how to use the ATMS for nonmonotonic reasoning [15], but the semantics of his inference relation does not meet the requirements we have identified for evaluating arguments, such as support for proof standards. Rather than trying to develop a new argument evaluation structure which meets our requirements on top of the ATMS, we have opted to develop a model of abduction sufficient for the purpose of goal selection on top of our existing Carneades argument evaluation structure.

The definition of relevance presented here continues our prior work on modeling issues [16,17,18], where issues are understood as relevant, contested statements. The details of these models vary considerably, as a consequence of their very different underlying formal models of argument. This line of work is not related to relevance logic, which aims to weaken classical logic to avoid claimed paradoxes of material and strict implication.

Assumption revision in our system is much simpler than general belief revision [19], since all of the formulas in a set of assumptions have been restricted to literals.

In [20], decision theory was used to select arguments to put forward in a formal argument game, making use of the expected utility of an argument, the probability of an argument being successful and the costs of the argument. Although this work is quite different from ours, since our aim is to help agents to select goals to work on, by searching for information which can be used to construct arguments about the goal, not to select from amoung a fixed set of arguments, both have in common the goal of helping agents with strategic issues when making moves in argumentation processes and both use cost functions for this purpose. An issue for further research is whether expected utilities, as well as costs, could be useful in our context.

Computational complexity issues may be interesting to investigate. The problem of computing positions is presumably intractable, since it depends on the reduction of formulas to their minimal disjunctive normal form, which is itself in general an intractable problem. However it may be possible to reformulate our problem to make it less complex, by using the structure of argument graphs to incrementally construct labels in minimal disjunctive normal form, rather than reducing them to this normal form after they have been constructed.

The model of abduction presented here has been implemented in the latest version of our Carneades[4] argumentation software. Our next steps include the development of pilot applications, as part of an effort to validate the model.

References

[1] Horst W.J. Rittel and Melvin M. Webber. Dilemmas in a general theory of planning. *Policy Science*, 4:155–169, 1973.

[2] Ronald P. Loui. Process and policy: resource-bounded non-demonstrative reasoning. *Computational Intelligence*, 14:1–38, 1998.

[3] Douglas Walton. *The New Dialectic: Conversational Contexts of Argument*. University of Toronto Press, Toronto; Buffalo, 1998. 24 cm.

[4] Kathleen Freeman and Arthur M. Farley. A model of argumentation and its application to legal reasoning. *Artificial Intelligence and Law*, 4(3-4):163–197, 1996.

[5] Thomas F. Gordon, Henry Prakken, and Douglas Walton. The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10-11):875–896, 2007.

[6] Thomas F. Gordon and Douglas Walton. Proof burdens and standards. In Iyad Rahwan and Guillermo Simari, editors, *Argumentation in Artificial Intelligence*, pages 239–260. Springer-Verlag, Berlin, Germany, 2009.

[7] Adam Wyner, Tom van Engers, and Anthony Hunter. Working on the argument pipeline: Through flow issues between natural language argument, instantiated arguments, and argumentation frameworks. In

[4] http://carneades.berlios.de

Proceedings of the Workshop on Computational Models of Natural Argument, Lisbon, Portugal, August 2010. to appear.

[8] Trevor Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[9] Trevor J.M. Bench-Capon, Sylvie Doutre, and Paul E. Dunne. Audiences in argumentation frameworks. *Artificial Intelligence*, 171(42-71), 2007.

[10] Stefan Ballnat and Thomas F. Gordon. Goal selection in argumentation processes – a formal model of abduction in argument evaluation structures. Technical report, Fraunhofer FOKUS, Berlin, Germany, 2010.

[11] Dan Sperber and Deirde Wilson. *Relevance: Communication and cognition*. Blackwell Publishers, 1995.

[12] Johan de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28(2):127–162, 1986.

[13] J. De Kleer. A general labeling algorithm for assumption-based truth maintenance. In *Proceedings of the 7th national conference on artificial intelligence*, pages 188–192, San Francisco, California, USA, 1988. Morgan Kaufmanns Publishers.

[14] Johann De Kleer. Problem solving with the ATMS. *Artificial Intelligence*, 28(2):197–224, 1986.

[15] Ulrich Junker. A correct non-monotonic ATMS. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 1049–1054. Morgan Kaufmann Publishers, Detroit, Michigan, 1989.

[16] Thomas F. Gordon. Issue spotting in a system for searching interpretation spaces. In *Proceedings of the Second International Conference on Artificial Intelligence and Law*, pages 157–164. Association for Computing Machinery (ACM), New York, 1989.

[17] Thomas F. Gordon. An abductive theory of legal issues. *International Journal of Man-Machine Studies*, 35:95–118, 1991.

[18] Thomas F. Gordon. *The Pleadings Game; An Artificial Intelligence Model of Procedural Justice*. Springer, New York, 1995. Book version of 1993 Ph.D. Thesis; University of Darmstadt.

[19] Carlos E. Alchourron, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, June 1985.

[20] Régis Riverert, Henry Prakken, Antonio Rotolo, and Giovanni Sartor. Heuristics in argumentation: A game-theoretical investigation. In Philippe Besnard, Sylvie Doutre, and Anthony Hunter, editors, *Compuational Models of Argument, Proceedings of COMMA 2008*, pages 324–335, 2008.