Propositional Argumentation and Causal Reasoning

Alexander Bochman
Computer Science Department,
Holon Academic Institute of Technology, Israel
e-mail: bochmana@hait.ac.il

Abstract

The paper introduces a number of propositional argumentation systems obtained by gradually extending the underlying language and associated monotonic logics. An assumption-based argumentation framework [Bondarenko *et al.*, 1997] will constitute a special case of this construction. In addition, a stronger argumentation system in a full classical language will be shown to be equivalent to a system of causal reasoning [Giunchiglia *et al.*, 2004]. The implications of this correspondence for the respective nonmonotonic theories of argumentation and causal reasoning are discussed.

1 Introduction

In this study we will demonstrate, among other things, that two relatively recent approaches to nonmonotonic reasoning, namely argumentation theory [Dung, 1995b; Bondarenko *et al.*, 1997] and a theory of causal nonmonotonic reasoning [McCain and Turner, 1997; Giunchiglia *et al.*, 2004; Bochman, 2003b; 2004a] are actually 'two faces of the same coin'. The common basis of these formalisms is a notion of explanation, which takes the form of an attack relation in the argumentation theory, and a causal relation in the theory of causal reasoning. It will be shown that, at some level of generalization, these two relations will be directly interdefinable.

Both the argumentation theory and theory of causal nonmonotonic reasoning have proved to be powerful frameworks for representing different kinds of nonmonotonic reasoning, including traditional nonmonotonic logics, logic programming, abduction and reasoning about actions and change. The present study should hopefully provide a basis for clarifying the representation capabilities of these formalisms, and a step toward a unified theory of nonmonotonic reasoning.

Another objective of this study consists in a systematic development of a *propositional* approach to argumentation, in which arguments are represented as (special kinds of) propositions. This approach will naturally lead us to developing the underlying *logics* of argumentation (cf. [Chesñevar *et al.*, 2000; Prakken and Vreeswijk, 2001]) and their associated semantics. As will be shown, the resulting argumentation systems will directly represent a number of important nonmonotonic formalisms.

The plan of the paper is as follows. We introduce first an abstract argumentation theory, and describe its four-valued semantics based on acceptance and rejection of arguments. Adding a 'global' negation to this language will give us a formalism that generalizes the assumption-based argumentation framework from [Bondarenko *et al.*, 1997]. Then we describe an alternative extension of the argumentation theory to the full classical propositional language. A special case of the latter formalism will be shown to be equivalent to a system of causal reasoning. Finally, we will consider the correspondence between argumentation and causal reasoning on the level of their associated nonmonotonic semantics.

2 Collective argumentation

As a general formal basis of argumentation theory, we will adopt the formalism of collective argumentation suggested in [Bochman, 2003a] as a 'disjunctive' generalization of Dung's argumentation theory. In this formalism, a primitive attack relation holds between *sets* of arguments: in the notation introduced below, $a \hookrightarrow b$ says that a set a of arguments attacks a set of arguments b. This fact implies, of course, that these two sets arguments are incompatible. $a \hookrightarrow b$ says, however, more than that, namely that the set of arguments a, being accepted, provides a reason, or explanation, for rejection of the set of arguments b. Accordingly, the attack relation will not in general be symmetric, since acceptance of b need not give reasons for rejection of a. In addition, the attack relation is not reducible to attacks between individual arguments. For instance, we can disprove some conclusion jointly supported by a disputed set of arguments, though no particular argument in the set, taken alone, could be held responsible for this.

In what follows, $a, b, c, \dots$ will denote finite sets of arguments, while $u, v, w, \dots$ arbitrary such sets. We will use the same agreements for the attack relation as for usual consequence relations. Thus, $a, a_1 \hookrightarrow b, B$ will have the same meaning as $a \cup a_1 \hookrightarrow b \cup \{B\}$, *etc*.

Definition 2.1. Let $\mathcal{A}$ be a set of *arguments*. A (collective) *attack relation* is a relation $\hookrightarrow$ on finite sets of arguments satisfying the following postulate:

Monotonicity If $a \hookrightarrow b$, then $a, a_1 \hookrightarrow b, b_1$.

Though defined primarily on finite sets of arguments, the attack relation will be extended to arbitrary such sets by imposing the compactness requirement: for any $u, v \subseteq \mathcal{A}$, (**Compactness**) $u \hookrightarrow v$ if and only if $a \hookrightarrow b$, for some finite $a \subseteq u$, $b \subseteq v$.

The original Dung's argumentation theory [Dung, 1995b] can be seen as a special case of collective argumentation that satisfies additional properties (cf. [Kakas and Toni, 1999]):

Definition 2.2. An attack relation is *normal* if no set of arguments attacks $\emptyset$, and the following condition is satisfied:

(**Locality**) If $a \hookrightarrow b, b_1$, then either $a \hookrightarrow b$, or $a \hookrightarrow b_1$.

For a normal attack, $a \hookrightarrow b$ holds iff $a \hookrightarrow A$, for some $A \in b$. As was shown in [Bochman, 2003a], the resulting argumentation theory coincides with that given in [Dung, 1995a].

By an *argument theory* we will mean an arbitrary set of attacks $a \hookrightarrow b$ between sets of arguments. Any argument theory Δ generates a unique least attack relation that we will denote by $\hookrightarrow_\Delta$. The latter can be described directly as follows:

$$u \hookrightarrow_\Delta v \ \text{ iff } \ a \hookrightarrow b \in \Delta, \ \text{for some } a \subseteq u, b \subseteq v.$$

It can be easily verified that an attack relation is normal if and only if it is generated in this sense by an argument theory consisting only of attack rules of the form $a \hookrightarrow A$.

2.1 Four-valued semantics

Collective argumentation can be given a four-valued semantics that will be instructive in describing the meaning of the attack relation. This semantics stems from the following understanding of an attack $a \hookrightarrow b$:

> *If all arguments in a are accepted, then at least one of the arguments in b should be rejected.*

The argumentation theory does not impose, however, the classical constraints on acceptance and rejection of arguments, so an argument can be both accepted and rejected, or neither accepted, nor rejected. Such an understanding can be captured formally by assigning any argument a *subset* of the set $\{t, f\}$, where t denotes acceptance (truth), while f denotes rejection (falsity). This is nothing other than the well-known *Belnap's interpretation* of four-valued logic (see [Belnap, 1977]). As a result, collective argumentation acquires a natural four-valued semantics described below.

Definition 2.3. An attack $a \hookrightarrow b$ will be said to *hold* in a four-valued interpretation ν of arguments, if either $t \notin \nu(A)$, for some $A \in a$, or $f \in \nu(B)$, for some $B \in b$.

An interpretation ν will be called a *model* of an argument theory Δ if every attack from Δ holds in ν.

Since an attack relation can be seen as a special kind of an argument theory, the above definition determines also the notion of a model for an attack relation.

For a set I of four-valued interpretations, we will denote by $\hookrightarrow_I$ the set of all attacks that hold in each interpretation from I. Then the following result is actually a representation theorem showing that the four-valued semantics is adequate for collective argumentation (see [Bochman, 2003a]).

Theorem 2.1. *$\hookrightarrow$ is an attack relation iff it coincides with $\hookrightarrow_I$, for some set of four-valued interpretations I.*

2.2 Assumption-based argumentation

The notion of an argument is often taken as primitive in argumentation theory, which allows for a possibility of considering arguments that are not propositional in character (e.g., arguments as inference rules, or derivations). As has been shown in [Bondarenko *et al.*, 1997], however, a powerful version of argumentation theory can be obtained by identifying arguments with propositions of a special kind called *assumptions*[1]. This representation can be refined further to a full-fledged theory of *propositional* argumentation in a certain well-defined propositional language.

Let us extend the language of arguments with a negation connective $\sim$ having the following semantic interpretation:

$$\sim\! A \text{ is accepted iff } A \text{ is rejected}$$
$$\sim\! A \text{ is rejected iff } A \text{ is accepted.}$$

The connective $\sim$ will be called a *global negation*, since it switches the evaluation contexts between acceptance and rejection. An axiomatization of this connective in argumentation theory can be obtained by imposing the following rules on the attack relation (see [Bochman, 2003a]):

$$A \hookrightarrow \sim\! A \qquad \sim\! A \hookrightarrow A$$
$$\text{If } a \hookrightarrow A, b \text{ and } a, \sim\! A \hookrightarrow b, \text{ then } a \hookrightarrow b \qquad \text{(AN)}$$
$$\text{If } a, A \hookrightarrow b \text{ and } a \hookrightarrow b, \sim\! A, \text{ then } a \hookrightarrow b$$

Attack relations satisfying the above postulates will be called *N-attack relations*. It turns out that the latter are interdefinable with certain consequence relations.

A *Belnap consequence relation* in a propositional language with a global negation $\sim$ is a Scott (multiple-conclusion) consequence relation $\Vdash$ satisfying the postulates

(Reflexivity) $A \Vdash A$;

(Monotonicity) If $a \Vdash b$ and $a \subseteq a'$, $b \subseteq b'$, then $a' \Vdash b'$;

(Cut) If $a \Vdash b, A$ and $a, A \Vdash b$, then $a \Vdash b$,

as well as the following Double Negation rules for $\sim$:

$$A \Vdash \sim\sim\! A \qquad \sim\sim\! A \Vdash A.$$

For any set u of propositions, we will denote by $\sim\! u$ the set $\{\sim\! A \mid A \in u\}$. Now, for a given N-attack relation, we can define the following consequence relation:

$$a \Vdash b \equiv \ a \hookrightarrow \sim\! b \qquad \text{(CA)}$$

Similarly, for any Belnap consequence relation we can define the corresponding attack relation as follows:

$$a \hookrightarrow b \equiv \ a \Vdash \sim\! b \qquad \text{(AC)}$$

As has been shown in [Bochman, 2003a], the above definitions establish an exact equivalence between N-attack relations and Belnap consequence relations. This correspondence allows us to represent an assumption-based argumentation framework from [Bondarenko *et al.*, 1997] entirely in the framework of attack relations.

Slightly changing the definitions from [Bondarenko *et al.*, 1997], an assumption-based argumentation framework can

[1] See also [Kowalski and Toni, 1996].

defined as a triple consisting of an underlying deductive system (including the current set of beliefs), a distinguished subset of propositions Ab called *assumptions*, and a mapping from Ab to the set of all propositions of the language that determines the *contrary* $\overline{A}$ of any assumption A.

Now, the underlying deductive system can be expressed directly in the framework of N-attack relations by identifying deductive rules $a \vdash A$ with attacks of the form $a \hookrightarrow \sim A$. Furthermore, the global negation $\sim$ can serve as a faithful logical formalization of the operation of taking the contrary. More precisely, given an arbitrary language $\mathcal{L}$ that does not contain $\sim$, we can *define* assumptions as propositions of the form $\sim A$, where $A \in \mathcal{L}$. Then, since $\sim$ satisfies double negation, a negation of an assumption will be a proposition from $\mathcal{L}$. Moreover, such a 'negative' representation of assumptions will agree with the applications of the argumentation theory to other nonmonotonic formalisms, described in [Bondarenko *et al.*, 1997]. Accordingly, N-attack relations can be seen as a proper generalization of the assumption-based framework.

3 Classical propositional argumentation

Taking seriously the idea of propositional argumentation, it is only natural to make a further step and extend the underlying language of arguments to a full classical propositional language. The latter step should be coordinated, however, with the inherently four-valued nature of an attack relation. And the way to do this amounts to requiring that the relevant classical connectives should behave in a usual classical way with respect to both acceptance and rejection of arguments.

As a first connective of this kind, we introduce the *conjunction* connective $\wedge$ on arguments that is determined by the following familiar semantic conditions:

$$A \wedge B \text{ is accepted iff } A \text{ is accepted and } B \text{ is accepted}$$
$$A \wedge B \text{ is rejected iff } A \text{ is rejected or } B \text{ is rejected}$$

As can be seen, $\wedge$ behaves as an ordinary classical conjunction with respect to acceptance and rejection of arguments. On the other hand, it is a four-valued connective, since the above conditions determine a four-valued truth-table for conjunction in the Belnap's interpretation of four-valued logic (see [Belnap, 1977]). The following postulates provide a syntactic characterization of this connective for attack relations:

$$\begin{aligned} a, A \wedge B \hookrightarrow b \ \text{ iff } \ a, A, B \hookrightarrow b \\ a \hookrightarrow A \wedge B, b \ \text{ iff } \ a \hookrightarrow A, B, b \end{aligned} \qquad (A_\wedge)$$

Collective attack relations satisfying these postulates will be called *conjunctive*. The next result shows that they give a complete description of the four-valued conjunction.

Corollary 3.1. *An attack relation is conjunctive if and only if it coincides with $\hookrightarrow_I$, for some set of four-valued interpretations I in a language with a four-valued conjunction $\wedge$.*

An immediate benefit of introducing conjunction into the language of argumentation is that any finite set of arguments a becomes reducible to a single argument $\bigwedge a$:

$$a \hookrightarrow b \quad \text{iff} \quad \bigwedge a \hookrightarrow \bigwedge b.$$

As a result, the collective attack relation in this language is reducible to an attack relation between individual arguments, as suggested already in [Dung, 1995b].

Having a conjunction at our disposal, we only have to add a classical negation $\neg$ in order to obtain a classical language. Moreover, since sets of arguments are reducible to their conjunctions, we can represent the resulting argumentation theory using just a binary attack relation on classical formulas.

As a basic condition on argumentation in the classical propositional language, we will require only that the attack relation should respect the classical entailment $\vDash$ in the precise sense of being monotonic with respect to $\vDash$ on both sides.

Definition 3.1. A *propositional attack relation* is a relation $\hookrightarrow$ on the set of classical propositions satisfying the following postulates:

(Left Strengthening) If $A \vDash B$ and $B \hookrightarrow C$, then $A \hookrightarrow C$;

(Right Strengthening) If $A \hookrightarrow B$ and $C \vDash B$, then $A \hookrightarrow C$;

(Truth) $\mathbf{t} \hookrightarrow \mathbf{f}$;

(Falsity) $\mathbf{f} \hookrightarrow \mathbf{t}$.

Left Strengthening says that logically stronger arguments should attack any argument that is attacked already by a logically weaker argument, and similarly for Right Strengthening. Truth and Falsity postulates characterize the limit cases of argumentation by stipulating that any tautological argument attacks any contradictory one, and vice versa.

There exists a simple definitional way of extending the above attack relation to a collective attack relation between arbitrary sets of propositions. Namely, for any sets u, v of propositions, we can define $u \hookrightarrow v$ as follows:

$$u \hookrightarrow v \ \equiv \ \bigwedge a \hookrightarrow \bigwedge b, \text{ for some finite } a \subseteq u \text{ and } b \subseteq v$$

The resulting attack relation will satisfy the properties of collective argumentation, as well as the postulates $(A_\wedge)$ for conjunction.

Finally, in order to acquire full expressive capabilities of an argumentation theory, we can add the global negation $\sim$ to the language. Actually, a rather simple characterization of the resulting collective argumentation theory can be obtained by accepting the postulates AN for $\sim$, plus the following rule that permits the use of classical entailment in attacks:

Classicality If $a \vDash A$, then $a \hookrightarrow \sim A$ and $\sim A \hookrightarrow a$.

It can be verified that the resulting system satisfies all the postulates for propositional argumentation. The system will be used later for a direct representation of default logic.

3.1 Semantics

A semantic interpretation of propositional attack relations can be obtained by generalizing four-valued interpretations to pairs (u, v) of deductively closed theories, where u is the set of accepted propositions, while v the set of propositions that are not rejected. Such pairs will be called *bimodels*, while a set of bimodels will be called a *binary semantics*.

Definition 3.2. An attack $A \hookrightarrow B$ will be said to be *valid* in a binary semantics $\mathcal{B}$ if there is no bimodel (u, v) from $\mathcal{B}$ such that $A \in u$ and $B \in v$.

We will denote by $\hookrightarrow_{\mathcal{B}}$ the set of attacks that are valid in a semantics $\mathcal{B}$. This set forms a propositional attack relation. Moreover, the following result shows that propositional attack relations are actually complete for the binary semantics.

Theorem 3.2. *$\hookrightarrow$ is a propositional attack relation if and only if it coincides with $\hookrightarrow_{\mathcal{B}}$, for some binary semantics $\mathcal{B}$.*

3.2 Probative and causal argumentation

We will introduce now some stronger propositional attack relations that will be shown to correspond to systems of causal inference. These stronger attack relations are obtained by adding the following quite reasonable postulates:

(Left Or) If $A \hookrightarrow C$ and $B \hookrightarrow C$, then $A \vee B \hookrightarrow C$;

(Right Or) If $A \hookrightarrow B$ and $A \hookrightarrow C$, then $A \hookrightarrow B \vee C$;

(Self-Defeat) If $A \hookrightarrow A$, then $\mathbf{t} \hookrightarrow A$.

Definition 3.3. A propositional attack relation will be called *probative* if it satisfies Left Or, *basic*, if it also satisfies Right Or, and *causal*, if it is basic and satisfies Self-Defeat.

Probative argumentation allows for reasoning by cases. Its semantic interpretation can be obtained by restricting bimodels to pairs (α, v), where α is a world (maximal classically consistent set). The corresponding binary semantics will also be called *probative*. Similarly, the semantics for basic argumentation is obtained by restricting bimodels to world pairs (α, β); such a binary semantics will be called *basic*. Finally, the *causal* binary semantics is obtained from the basic semantics by requiring further that (α, β) is a bimodel only if (β, β) is also a bimodel.

Corollary 3.3. *A propositional attack relation is probative [basic, causal] iff it is determined by a probative [resp. basic, causal] binary semantics.*

Basic propositional argumentation can already be given a purely four-valued semantic interpretation, in which the classical negation $\neg$ has the following semantic description[2]:

$$\neg A \text{ is accepted iff } A \text{ is not accepted}$$
$$\neg A \text{ is rejected iff } A \text{ is not rejected}$$

A syntactic characterization of this connective in collective argumentation can be obtained by imposing the rules

$$\begin{gathered} A, \neg A \hookrightarrow \qquad \hookrightarrow A, \neg A \\ \text{If } a, A \hookrightarrow b \text{ and } a, \neg A \hookrightarrow b \text{ then } a \hookrightarrow b \\ \text{If } a \hookrightarrow b, A \text{ and } a \hookrightarrow b, \neg A \text{ then } a \hookrightarrow b \end{gathered} \qquad (A_\neg)$$

Then a basic propositional attack relation can be alternatively described as a collective attack relation satisfying the rules $(A_\wedge)$ and $(A_\neg)$. Moreover, the global negation $\sim$ can be added to this system just by adding the corresponding postulates (AN). It turns out, however, that the global negation is *eliminable* in this setting via to the following reductions:

$$\begin{array}{ll} a, \sim A \hookrightarrow b \equiv a \hookrightarrow b, \neg A & a \hookrightarrow \sim A, b \equiv a, \neg A \hookrightarrow b \\ a, \neg\sim A \hookrightarrow b \equiv a \hookrightarrow b, A & a \hookrightarrow \neg\sim A, b \equiv a, A \hookrightarrow b \end{array} \qquad (R_\sim)$$

[2] This connective coincides with the local four-valued negation from [Bochman, 1998].

As a result, the basic attack relation can be safely restricted to an attack relation in a classical language.

Finally, the rule Self-Defeat of causal argumentation gives a formal representation for an often expressed desideratum that self-conflicting arguments should not participate in defeating other arguments (see, e.g., [Bondarenko *et al.*, 1997]). This aim is achieved in our setting by requiring that such arguments are attacked even by tautologies, and hence by any argument whatsoever.

3.3 Argumentation vs. causal inference

Probative attack relations will now be shown to be equivalent to general production inference relations from [Bochman, 2003b; 2004a], a variant of input-output logics from [Makinson and van der Torre, 2000].

A *production inference relation* is a relation $\Rightarrow$ on the set of classical propositions satisfying the following rules:

(Strengthening) If $A \vDash B$ and $B \Rightarrow C$, then $A \Rightarrow C$;

(Weakening) If $A \Rightarrow B$ and $B \vDash C$, then $A \Rightarrow C$;

(And) If $A \Rightarrow B$ and $A \Rightarrow C$, then $A \Rightarrow B \wedge C$;

(Truth) $\mathbf{t} \Rightarrow \mathbf{t}$;

(Falsity) $\mathbf{f} \Rightarrow \mathbf{f}$.

A production rule $A \Rightarrow B$ can be informally interpreted as saying that A *produces*, or *explains* B. A characteristic property of production inference is that reflexivity $A \Rightarrow A$ does not hold for it. Production rules are extended to rules with sets of propositions in premises by requiring that $u \Rightarrow A$ holds for a set u of propositions iff $\bigwedge a \Rightarrow A$, for some finite $a \subseteq u$. $\mathcal{C}(u)$ will denote the set of propositions produced by u:

$$\mathcal{C}(u) = \{A \mid u \Rightarrow A\}$$

The production operator $\mathcal{C}$ plays much the same role as the usual derivability operator for consequence relations.

A production inference relation is called *basic*, if it satisfies

(Or) If $A \Rightarrow C$ and $B \Rightarrow C$, then $A \vee B \Rightarrow C$.

and *causal*, if it is basic and satisfies, in addition

(Coherence) If $A \Rightarrow \neg A$, then $A \Rightarrow \mathbf{f}$.

It has been shown in [Bochman, 2003b] that causal inference relations provide a complete description of the underlying logic of causal theories from [McCain and Turner, 1997] (see also [Giunchiglia *et al.*, 2004]).

It turns out that the binary semantics, introduced earlier, is appropriate also for interpreting production inference:

Definition 3.4. A rule $A \Rightarrow B$ is *valid* in a binary semantics $\mathcal{B}$ if, for any bimodel $(u, v) \in \mathcal{B}$, $A \in u$ only if $B \in v$.

As has been shown in [Bochman, 2004a], the above semantics is adequate for production inference relations. Moreover, the semantics for basic production inference can be obtained by restricting bimodels to world pairs (α, β), while the semantics for causal inference is obtained by requiring, in addition, that (α, β) is a bimodel only if (α, α) is also a bimodel.

Now, the correspondence between probative argumentation and production inference can be established directly on the syntactic level using the following definitions:

$$A \Rightarrow B \;\equiv\; \neg B \hookrightarrow A; \qquad \text{(PA)}$$

$$A \hookrightarrow B \;\equiv\; B \Rightarrow \neg A. \qquad \text{(AP)}$$

Under these correspondences, the rules of a probative attack relation correspond precisely to the postulates for production relations. Moreover, the correspondence extends also to a correspondence between basic and causal argumentation, on the one hand, and basic and causal production inference, on the other. Hence the following result is straightforward.

Lemma 3.4. *If $\hookrightarrow$ is a probative [basic, causal] attack relation, then (PA) determines a [basic, causal] production inference relation, and vice versa, if $\Rightarrow$ is a [basic, causal] production inference relation, then (AP) determines a probative [basic, causal] attack relation.*

Remark. A seemingly more natural correspondence between propositional argumentation and production inference can be obtained using the following definitions:

$$A \Rightarrow B \;\equiv\; A \hookrightarrow \neg B \qquad A \hookrightarrow B \;\equiv\; A \Rightarrow \neg B.$$

By these definitions, A explains B if it attacks $\neg B$, and vice versa, A attacks B if it explains $\neg B$. Unfortunately, this correspondence, though plausible by itself, does not take into account the intended understanding of arguments as (negative) assumptions. As a result, it cannot be extended directly to the correspondence between the associated nonmonotonic semantics, described below.

4 Nonmonotonic semantics

The correspondence between argumentation and causal reasoning has been established above on the level of underlying logical (monotonic) formalisms. In this section we will describe how this correspondence can be extended to the associated nonmonotonic semantics.

In [Bondarenko *et al.*, 1997], the assumption-based argumentation framework has been *instantiated* to capture existing nonmonotonic formalisms. In other words, it has been shown how particular nonmonotonic systems can be viewed as assumption-based frameworks just by defining assumptions and their contraries. We will show, however, that the propositional argumentation theory allows us to give a direct representation of Reiter's default logic [Reiter, 1980].

Given a system of propositional argumentation in the classical language augmented with the global negation $\sim$, we will interpret Reiter's default rule $a{:}b/A$ as an attack

$$a, \sim\neg b \hookrightarrow \sim A,$$

or, equivalently, as a rule $a, \sim\neg b \Vdash A$ of the associated Belnap consequence relation. Similarly, an axiom A of a default theory will be interpreted as an attack $\mathbf{t} \hookrightarrow \sim A$. For a default theory Δ, we will denote by $tr(\Delta)$ the corresponding argument theory obtained by this translation.

By our general agreement, by *assumptions* we will mean propositions of the form $\sim A$, where A is a classical proposition. For a set u of classical propositions, we will denote by $\tilde{u}$ the set of assumptions $\{\sim A \mid A \notin u\}$. Finally, a set w of assumptions will be called *stable* in an argument theory Δ if, for any assumption A, $A \in w$ iff $w \not\hookrightarrow_\Delta A$. Then we have

Theorem 4.1. *A set u of classical propositions is an extension of a default theory Δ if and only if $\tilde{u}$ is a stable set of assumptions in $tr(\Delta)$.*

The above result is similar to the corresponding representation result in [Bondarenko *et al.*, 1997] (Theorem 3.10), but it is much simpler, and is formulated entirely in the framework of propositional attack relations. The simpler representation was made possible due to the fact that propositional attack relations already embody the deductive capabilities treated as an additional ingredient in assumption-based frameworks.

As our next result, we will establish a correspondence between the nonmonotonic semantics of causal inference relations and that of causal argumentation.

The nonmonotonic semantics of a causal inference relation is a set of its *exact worlds*, namely worlds α such that $\alpha = \mathcal{C}(\alpha)$ (see [Bochman, 2004a]). Such a world satisfies the rules of the causal relation, and any proposition that holds in it is explained by the causal rules.

A *causal theory* is an arbitrary set of production rules. By a nonmonotonic semantics of a causal theory Δ we will mean the exact worlds of the least causal relation containing Δ.

The correspondence between exact worlds and stable sets of assumptions is established in the next theorem.

Theorem 4.2. *If Δ is a causal theory, and Δ_a its corresponding argument theory given by (AP), then a world α is an exact world of Δ iff $\tilde{\alpha}$ is a stable set of assumptions in Δ_a.*

The above result shows, in effect, that propositional argumentation subsumes causal reasoning as a special case. Moreover, it can be shown that causal attack relations constitute a strongest argumentation system suitable for this kind of nonmonotonic semantics.

As an application of the above correspondence, we will describe now an alternative argumentation-based representation of logic programming.

A *general logic program* Π is a set of rules of the form

$$\mathbf{not}\, d, c \leftarrow a, \mathbf{not}\, b \tag{*}$$

where a, b, c, d are sets of propositional atoms. These are program rules of a most general kind that contain disjunctions and negations as failure $\mathbf{not}$ in their heads. As has been shown in [Bochman, 2004b], general logic programs are representable as causal theories obtained by translating program rules (*) as causal rules

$$d, \neg b \Rightarrow \bigwedge a \to \bigvee c,$$

and adding a formalization of the Closed World Assumption:

(Default Negation) $\neg p \Rightarrow \neg p$, for any propositional atom p.

Now, due to the correspondence between causal reasoning and argumentation, this causal theory can be transformed (using (PA)) into an argument theory that consists of attacks

$$a, \neg c \hookrightarrow \neg b, d \tag{AL}$$

plus the 'argumentative' Closed World Assumption:

(Default Assumption) $p \hookrightarrow \neg p$, for any atom p.

Let $tr(\Pi)$ denote the argument theory obtained by this translation from a logic program Π. Then we obtain

Theorem 4.3. *A set u of propositional atoms is a stable model of a logic program Π iff $\tilde{u}$ is a stable set of assumptions in $tr(\Pi)$.*

It is interesting to note that, due to the reduction rules ($R_\sim$) for the global negation $\sim$, described earlier, the above representation (AL) of program rules is equivalent to $a, \sim b \hookrightarrow \sim c, d$, and therefore to the inference rules

$$a, \sim b \Vdash c, \sim d$$

of the associated Belnap consequence relation. For normal logic programs, this latter representation coincides with that given in [Bondarenko *et al.*, 1997].

5 Conclusions

Two main objectives have been pursued in this study. The first consisted in showing that propositional argumentation suggests a viable and useful extension of an abstract argumentation theory that allows us to endow argumentation with full-fledged logical capabilities. It has been shown, in particular, that propositional argumentation subsumes assumption-based frameworks, and provides a direct representation of default logic, causal reasoning and logic programming. It could be expected that further development of this propositional approach to argumentation may bring us additional theoretical and practical benefits.

The second aim was to demonstrate that causal reasoning can be seen as an important kind of argumentation. It has been shown in this respect that causal inference relations and their semantics exactly correspond to a special, quite strong, kind of propositional attack relations and their associated stable semantics. This correspondence established a basic link between argumentation and causal reasoning, and it can be extended in both directions. To begin with, the argumentation theory has suggested a number of weaker semantic models, such as admissible sets and complete extensions (see [Bondarenko *et al.*, 1997; Dung, 1995a]), and it seems worth to inquire whether such models correspond to reasonable semantics for causal reasoning. On the other hand, a number of alternative models for causal reasoning has been suggested in [Geffner, 1992; 1997] (see also [Darwiche and Pearl, 1994]), and here it seems plausible to suppose that the correspondence between causal reasoning and argumentation could be helpful in analyzing and evaluating such models. These are, however, the topics for further study.

References

[Belnap, 1977] N. D. Belnap, Jr. A useful four-valued logic. In M. Dunn and G. Epstein, editors, *Modern Uses of Multiple-Valued Logic*, pages 8–41. D. Reidel, 1977.

[Bochman, 1998] A. Bochman. Biconsequence relations: A four-valued formalism of reasoning with inconsistency and incompleteness. *Notre Dame Journal of Formal Logic*, 39(1):47–73, 1998.

[Bochman, 2003a] A. Bochman. Collective argumentation and disjunctive logic programming. *Journal of Logic and Computation*, 9:55–56, 2003.

[Bochman, 2003b] A. Bochman. A logic for causal reasoning. In *Proceedings IJCAI'03*, Acapulco, 2003. Morgan Kaufmann.

[Bochman, 2004a] A. Bochman. A causal approach to nonmonotonic reasoning. *Artificial Intelligence*, 160:105–143, 2004.

[Bochman, 2004b] A. Bochman. A causal logic of logic programming. In D. Dubois, C. Welty, and M.-A. Williams, editors, *Proc. Ninth Conference on Principles of Knowledge Representation and Reasoning, KR'04*, pages 427–437, Whistler, 2004.

[Bondarenko *et al.*, 1997] A. Bondarenko, P. M. Dung, R. A. Kowalski, and F. Toni. An abstract, argumentation-theoretic framework for default reasoning. *Artificial Intelligence*, 93:63–101, 1997.

[Chesñevar *et al.*, 2000] C. I. Chesñevar, A. G. Marguitman, and R. P. Loui. Logical models of argument. *ACM Computing Surveys*, 32:337–383, 2000.

[Darwiche and Pearl, 1994] A. Darwiche and J. Pearl. Symbolic causal networks. In *Proceedings AAAI'94*, pages 238–244, 1994.

[Dung, 1995a] P. M. Dung. An argumentation-theoretic foundation for logic programming. *J. of Logic Programming*, 22:151–177, 1995.

[Dung, 1995b] P. M. Dung. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-persons games. *Artificial Intelligence*, 76:321–358, 1995.

[Geffner, 1992] H. Geffner. *Default Reasoning. Causal and Conditional Theories*. MIT Press, 1992.

[Geffner, 1997] H. Geffner. Causality, constraints and the indirect effects of actions. In *Proceedings Int. Joint Conf. on Artificial Intelligence, IJCAI'97*, pages 555–561, 1997.

[Giunchiglia *et al.*, 2004] E. Giunchiglia, J. Lee, V. Lifschitz, N. McCain, and H. Turner. Nonmonotonic causal theories. *Artificial Intelligence*, 153:49–104, 2004.

[Kakas and Toni, 1999] A. C. Kakas and F. Toni. Computing argumentation in logic programming. *Journal of Logic and Computation*, 9:515–562, 1999.

[Kowalski and Toni, 1996] R. A. Kowalski and F. Toni. Abstract argumentation. *Artificial Intelligence and Law*, 4:275–296, 1996.

[Makinson and van der Torre, 2000] D. Makinson and L. van der Torre. Input/Output logics. *Journal of Philosophical Logic*, 29:383–408, 2000.

[McCain and Turner, 1997] N. McCain and H. Turner. Causal theories of action and change. In *Proceedings AAAI-97*, pages 460–465, 1997.

[Prakken and Vreeswijk, 2001] H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. In D. M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 219–318. Kluwer, 2nd edition, 2001.

[Reiter, 1980] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.