Using Abductive Logic Agents for Modelling Judicial Evaluation of Criminal Evidence

A. Ciampolini and P. Torroni
Dipartimento di Elettronica, Informatica e Sistemistica
Università di Bologna
Viale Risorgimento 2, 40136 Bologna, Italy
{aciampolini, ptorroni}@deis.unibo.it

Abstract. In this paper we present a novel approach to the modelling of judicial evaluation of criminal evidence, an approach based on an abductive multi-agent system. Legal justification in such a context has the main objective of finding the relatively most plausible explanations for some given pieces of evidence, especially (but not necessarily) in the trial of a criminal case. In our approach, the process of justification is done through the collaborative abductive reasoning of agents, which operate within a logic-based architecture called *ALIAS*.This allows us the modular composition of the knowledge of different *dramatis personæ* involved in the trial: the detective, witnesses, suspects, judges, and so forth. Having represented each actor in the trial by a single *ALIAS* agent, we are able to dynamically group and coordinate them for the explanation of goals (such as, for instance, pieces of evidence or any given observation).

We tested the proposed approach on the Peyer case, which was tried in California; it is an example borrowed from the literature that has been adopted as a test bed by previous abduction-based approaches. We will show that using ALIAS agents in legal justification allows not only to produce plausible explanations for observed pieces of evidence, but also to detect either collusions or inconsistencies among trial characters. Moreover, we will show how legal justification with the proposed approach could take into consideration also the credibility of persons (e.g., witnesses) involved in the trial.

1 Introduction

In recent years, the interest for intelligent agents has considerably grown from both the theoretical and the practical point of view [Gini and Ishida, 2002]. Notwithstanding disagreement on what is an “agent” (see Section 2.3 below), the agent paradigm, in fact, is well suited for representing applications that merge features inherited from both the distributed systems area (such as locality, distribution, interaction, mobility etc.) and artificial intelligence (with features such as reasoning, adaptation, etc.). In particular, intelligent agents require both reasoning capabilities and social abilities, which make interaction and coordination among agents possible.

Abduction has been widely recognized as a powerful mechanism for hypothetical reasoning in presence of incomplete knowledge [Cox and Pietrzykowski, 1986, Eshgi and Kowalski, 1989, Kakas and Mancarella, 1990a]: it is generally understood as reasoning from effects to causes, and also captures other important issues such as reasoning with defaults and beliefs (see for instance [Kowalski, 1990, Poole, 1988]). For these reasons, abduction has been already – in a few instances conspicuous in the literature, if not extensively – applied to legal reasoning, especially as an explanatory inference technique in legal trials [Thagard, 1989, Fox and Josephson, 1990].

In this paper we present a novel approach for legal justification of verdicts in criminal trials, based on an abductive multi-agent system. Legal evaluation of evidence has the main objective of finding the cogent explanations for some given pieces of evidence or observations, and unlike civil cases, in criminal trials the utility function considers it prohibitive that an innocent defendant be mistakenly convicted. The process of legal evidentiary reasoning is arguably well served by abductive reasoning (even though, especially in the U.S., rules intervene that make some kinds of evidence inadmissible, which by abduction alone would not have been excluded at all; their rationale is that human rights take precedence over getting a better grasp of the events being reconstructed. Legal truth is *not* the same as factual truth).

When trying to model, in AI, legal evidentiary reasoning, abduction – which more generally could be particularly useful when the knowledge base representing the problem domain is either incomplete or multiple – is appropriate, as events of a crime being tried need be reconstructed, and some pieces of knowledge about the crime are often missing; moreover, multiple knowledge is available in the case of several witnesses, when it could happen that some of them provide a different version about the crime, and some inconsistencies could arise.

We deal with legal evaluation of evidence by means of ALIAS, an agent architecture where agents are logic-based and capable to reason with abduction. ALIAS agents can be dynamically grouped and their reasoning can be coordinated by means of proper operators: namely, *collaboration* and *competition*.

The multi-agent approach we are proposing relies on several agents, each representing a distinct actor in the trial, e.g., the detective, suspects, witnesses, and so forth. Once we have represented each actor by a single abductive agent, we are able to dynamically group (some of) the agents and coordinate them for the explanation of goals. As for traditional abductive reasoning, the system starts from a set S of observed pieces of evidence, to provide one or more sets of hypotheses as possible explanations for S, using the knowledge of a group of agents. In addition, the dynamic composition of knowledge bases of agents allows to perform some additional tasks that could help both the detectives in police investigations and judicial factfinders (be these lay, as in trials by jury, or trained judges, as in bench trials) in solving the case. For instance, it is (hopefully) possible to check the coherence of distinct witnesses, thus possibly to detect either collusion or false testimony, or the unintentionally inaccurate testimony, such as owing to *confabulation*, when the witnesses integrate a narrative inferentially instead of perceptually.

It is worthwhile noticing that, in realistic cases, finding explanations for a given evidence could be unsuccessful due to various reasons: for instance, not all the available pieces of evidence could be reliable, depending, for instance, on the source. We faced this problem by representing the reliability either of each *dramatis persona* or even of single pieces of knowledge (e.g. a passage in a witness interrogation) by means of proper hypotheses. In this way we can obtain justifications labelled with reliability hypotheses, that could further help detectives and factfinders in understanding behaviours and roles of actors involved in the trial. We tested the system on the Peyer case, a well known trial that has been adopted as a testbed by previous abduction-based approaches [Thagard, 1989, Fox and Josephson, 1990, Punch III et al., 1990].

The paper is organized ad follows. Section 2 gives some preliminary notions on abduction and presents *ALIAS*. In section 3 we show how to apply abduction and agents to the Peyer case, and give some general guidelines to be followed when using abductive agents in legal trials. A discussion of related literature is presented in section 4. Conclusions follow.

2 Logic-based abductive reasoning and Abductive Logic Agents

In this section we give some preliminary notions on abductive logic programming and multi-agent systems. In particular, we firstly introduce logic-based abduction, as a tool for hypothetical reasoning. Next, we present *ALIAS*, the multi-agent architecture based on abductive logic agents that we use in the following to model the process of legal justification.

2.1 Abductive logic programming

An assumption that is often made in multi-agent systems is that individuals only have a partial knowledge of the environment they are situated in. This is particularly true in the legal application domains, if we want to use agents to model physical persons or different jurisdictional institutions, as we will see in the sequel. Therefore, in most knowledge-intensive applications of agents, it is often the case that an intelligent agent requires some sort of *guess* about a computation which cannot be performed locally, since the local knowledge is incomplete. In the legal setting, the incomplete knowledge could be, for instance, the information that each person involved in a legal trial (either the detective, a witness or a judge or a juror) has about the crime. In a logic programming perspective, we could rephrase it by saying that a guess is needed because the agent has a *goal* which cannot be solved locally. In this situation, the Closed World Assumption [Clark, 1978] usually adopted in logic programming can be no longer assumed, and some form of *open* or *abductive* reasoning has to be considered. Different forms of abduction, though not in logic programming, have already been used for legal justification by different authors [Thagard, 1989, Fox and Josephson, 1990].

Abduction is generally understood as reasoning from effects to causes, and captures other important issues, such as reasoning with defaults and beliefs (see for instance [Kowalski, 1990, Poole, 1988]). Incomplete knowledge is handled by labeling some pieces of information as abducibles, i.e., possible hypotheses which can be assumed, provided that they are consistent with the current knowledge base. In the context of intelligent agents, abduction can be regarded as a way of *dynamically* enlarging the agent's knowledge with abducible atoms. In the case of a single agent that uses a proof to reason abductively, such proof will guarantee that the abduced hypotheses that enlarge the agent's knowledge base are consistent with its knowledge. For this purpose, an agent's knowledge base is equipped with some integrity constraints, that must never be violated.

Before giving a formal definition of abductive logic programs, we will briefly recall some basic concepts about logic programs.

A logic program is a set of clauses [1]

$$A \leftarrow L_1, \ldots, L_n$$

[1] From now on, we will use logic programming notation as it is defined in [Lloyd, 1987]. Moreover, in this work we will adopt the Edinburgh notation for standard Prolog, where variable names start with upper case or '_', and constants start with lower case.

where A is an atom, and each L_i is a literal, i.e., an atom B or its negation $not\ B$. An *atom* is an n-ary predicate symbol followed by a list of n terms $p(t_1, \ldots, t_n)$, or `true`, or `false`. A *literal* is an atom or the negation of an atom. A *term* is a constant, a variable or a compound term.

Note that `not` represents default negation or negation as failure. All variables in $A, L_1, \ldots, L_n$ are assumed to be universally quantified from the outside. A is referred to as the *head* of the clause and $L_1, \ldots, L_n$ as its *body*.

Definition 1 (*abductive logic program*) An *abductive logic program* (ALP) is a triple $\langle P, \mathcal{A}, IC\rangle$ where P is a logic program, possibly with abducible atoms in clause bodies; $\mathcal{A}$ is a set of *abducible predicates*, i.e. *open* predicates which can be used to form explaining sentences; IC is a set of integrity constraints.

Given an abductive program $\langle P, \mathcal{A}, IC\rangle$ and a sentence G, the purpose of abduction is to find a (possibly minimal) set of abducible atoms δ which together with P "entails" G, and such that $P \cup \delta$ "satisfies" IC. Depending on the chosen semantics for the given logic program [Kakas et al., 1998], various notions of "entailment" are possible, such that, for every ground literal in the language, either the literal or its complement is entailed, and such that no literal and its complement are entailed at the same time. In the following, we will use the following notation for the concept of "abductive entailment":

$$\langle P, \mathcal{A}, IC\rangle \overset{abd}{\vdash}_{\delta} \; G$$

and it is possible to formalize its declarative semantics in the following way:

Definition 2 (*declarative semantics of an ALP*) Let $\langle P, \mathcal{A}, IC\rangle$ be an abductive logic program. Then, given a goal G, the *declarative semantics* of $\langle P, \mathcal{A}, IC\rangle \overset{abd}{\vdash}_{\delta} \; G$ is defined as follows:

$$\begin{cases} P \cup \delta \models G \\ P \cup \delta \;\; \text{“\textit{satisfies}”} \;\; IC \\ \delta \subseteq \mathcal{A} \end{cases}$$

where $\models$ is the entailment symbol.

There are also different notions of "integrity constraint satisfaction" in literature. Different procedures / semantics will adopt different such notions. Moreover, there is not a unique *syntax* to write integrity constraints: depending on the adopted proof-procedure, they could be written for example as denials, as in the case of the Kakas-Mancarella proof-procedure (see Section 2.2 below), or as forward rules (implications), as in the IFF proof-procedure. All proofs require that each integrity constraint contains at least one abducible, and that abducible predicates are not defined, i.e., there is no clause in an ALP that has an abducible in the head. In fact, this would mean that the abducible could be derived by deduction, while we consider abducibles as open predicates that can be assumed true or false, depending on the goal to solve. It is worthwhile noticing that, given a program P, possibly containing definitions of abducible predicates, it is possible to syntactically transform P into another semantically equivalent program, by means of auxiliary predicates, where no abducible is defined. Such transformation is shown in [Kakas et al., 1998]. The following example shows a simple abductive logic program, written according to the syntax of integrity constraints required by the

Kakas-Mancarella proof-procedure.

Example 1 Let us consider the following program (inspired by the Peyer case, described in section 3), expressing two possible explanations for the retrieval of the body of Cara Knott, in the Peyer trial (in which "Peyer" is the name of the main defendant in the trial). In this example the atoms $peyer\_killed\_knott$ and $someone\_else\_killed\_knott$ are abducible predicates:

found_knott's_body $\leftarrow$ *peyer_killed_knott.*
found_knott's_body $\leftarrow$ *someone_else_killed_knott.*

with integrity constraint:

$\leftarrow$ *someone_else_killed_knott, peyer_killed_knott.*

In order to satisfy such constraint, the predicates *someone_else_killed_knott* and *peyer_killed_knott* cannot be assumed both true at the same time. The observation *found_knott's_body* is therefore explained by two (minimal) sets of sentences, δ_1 and δ_2:

$\delta_1 = \{$*someone_else_killed_knott, not peyer_killed_knott*$\}$ and
$\delta_2 = \{$*peyer_killed_knott, not someone_else_killed_knott*$\}$

According to [Eshgi and Kowalski, 1989], negation as default, possibly occurring in clause bodies, can be recovered into abduction by replacing negated literals of the form $not\ a$ with a new positive, abducible atom not_a and by adding the integrity constraint $\leftarrow a, not\_a$ to IC. The natural syntactic correspondence between a standard atom and its negation by default is given by the following notion of complement:

$$\bar{l} = \begin{cases} \alpha & if \;\; l = not\_\alpha \\ not\_\alpha & otherwise \end{cases}$$

where α is an atom. All negated literals in an ALP are considered abducibles.

The integrity constraints used for handling negation as default, like the constraint $\leftarrow p, not\ p$, are considered trivial, and we will leave them implicit and suppose that they are replicated in each agent's knowledge base.

We will briefly introduce now the Kakas-Mancarella proof-procedure for abductive derivation, that implements the declarative semantics defined above. This is not the only abductive proof-procedures defined in the logic programming literature [Fung and Kowalski, 1997, Denecker and Schreye, 1998, Console et al., 1991], but it is the one adopted by ALIAS agents.

2.2 The Kakas-Mancarella proof-procedure

The Kakas-Mancarella (KM, for short) proof-procedure [Kakas and Mancarella, 1990b] is an extension of the proof-procedure for ordinary logic programming proposed by [Eshgi and Kowalski, 1989]. Both procedures extend the basic resolution mechanism adopted in SLD and SLDNF in ordinary logic programming by introducing the notion of *abductive* and *consistency* derivations.

Intuitively, an *abductive* derivation is the usual logic programming derivation suitably extended in order to consider abducibles. When an abducible atom h is encountered during this derivation, it is assumed, provided this is consistent. The consistency check of a hypothesis, then, starts the second kind of derivation. The *consistency* derivation verifies that, when the hypothesis h is assumed and added to the current set of hypotheses, any integrity constraint containing h fails (i.e., the bodies of all the integrity constraints containing h are false). During this latter procedure, when an abducible L is encountered, in order to prove its failure, an abductive derivation for its complement, $\overline{L}$, is attempted. In other words, as we said before, an abductive derivation computes a set of hypotheses "entailing" (with respect to the given logic program) a given goal/query/observation, starting from a given set of initial hypotheses. In order to compute such set and guarantee that the integrity constraints are "satisfied", an abductive derivation might require subsidiary consistency derivations, which in turn might require additional abductive derivations to ensure integrity constraint "satisfaction".

The only detail that we would like to give about the KM is about the syntax of abductive logic programs. In its original formulation, the KM requires that P is a propositional logic program, expressed in form of ground clauses. The integrity constraints are expressed in the form of denials of atoms, as in Example 1; at least one of such atoms must be abducible. Integrity constraint satisfaction in KM requires that at least one atom in the body of an integrity constraint be false. For instance, given the constraint

$\leftarrow a, b, c.$

and the set $\{a, b\}$ of abducibles, if we want to assume a to be true, we must either assume b false (by abducing $not\ b$) or prove that c is false. Both ways may require a deductive derivation in P, and possibly some more consistency steps. Such process is iterated until we find a consistent and "stable" set of abducibles δ, in the sense that further consistency steps do not result in an enlargement of such δ. The KM proof-procedure can be adopted also with first order logic programs (not only propositional): in that case, all variables in the body of an integrity constraint are assumed to be universally quantified from the outside. The KM proof-procedure has been proven sound [Brogi et al., 1997], for some specific notions of "entailment" and "satisfaction", with respect to the semantics defined in [Dung, 1995] and weakly complete for the 3-valued stable model semantics of [Brogi et al., 1997], as well as the argumentation-semantics defined in [Toni, 1995].

2.3 Abductive Logic-based multi-agent system: the ALIAS architecture

The agent metaphor is widely used in distributed artificial intelligence, although no real agreement has ever been reached on the definition of agent. In fact, the term 'agent' has been used to describe a broad range of computational entities, ranging from simple systems to very large ones [Gini and Ishida, 2002]. In our work, we use the agent metaphor as a model of a real world legal case, and we specify the agent knowledge by using the logic programming paradigm [van Emden and Kowalski, 1976], and particularly, a logic language. We will define here the concept of agent system, from a logic-programming perspective.

Definition 3 (*agent system*) An *agent system* is a finite set A, where each $x \in A$ is a ground term, representing the name of an agent, equipped with a *knowledge base* $\mathcal{K}(x)$. Each $\mathcal{K}(x)$

includes an abductive logic program.

We assume that, in an agent system, the agents share a common language. In particular, in *ALIAS*, the individuals will have separate knowledge bases but will refer to a common set of abducibles $\mathcal{A}$.

In ALIAS, each agent may thus autonomously reason on its own knowledge base and find abductive explanations to submitted goals.

Furthermore, as far as conflicting agents beliefs and goals are concerned, we propose a technique aimed to maintaining consistency in an environment made of multiple individuals, and to aiding the agent reasoning through possible conflicts of beliefs. For this purpose, Abductive Logic Agents can dynamically join into groups, or *bunches*, with the purpose, for instance, of finding the solution of a given goal in a *collaborative* way. In this perspective, although the set of program clauses and integrity constraints might differ from agent to agent, the set of abducible predicates (default predicates included) is the same for all the agents in a bunch. This implies that when proving a given goal, if an agent A assumes a new hypothesis h, all the arguing agents (i.e., the agents belonging to the same bunch) must check the consistency of h with their own integrity constraints. These checks could raise new hypotheses, whose consistency within the bunch has to be recursively checked. Therefore, in ALIAS, the abductive explanation of a goal within a bunch of agents is a set of abduced hypotheses, agreed by all agents in the bunch.

For this purpose, the current *ALIAS* implementation provides mechanisms to support agent bunches, local abduction and global consistency checks. As far as local abduction is concerned, the current implementation adopts the KM proof-procedure.[2]

In order to guarantee consistency of abduced hypotheses within a bunch, we adopt a two-phase algorithm, which first collects the results of local[3] abduction processes by running the KM proof-procedure, and then, depending on the chosen coordination policy, checks their consistency, and, in case of success obtains a consistent set of hypotheses Δ. The following example will show how abduction is performed within a bunch of two agents.

Example 2 Let us consider again Example 1, where all the knowledge needed to find an explanation to the observation *found_knott's_body* is enclosed in a single agent, that we will call A_0. Now, let us suppose that another agent, say A_1, has a additional knowledge about the same problem. In particular, let A_1's knowledge base be:

peyer_has_alibi.

and the following integrity constraint:

$\leftarrow$ *peyer_killed_knott, peyer_has_alibi.*

In order for them to communicate the result of their reasoning, expressed in form of conjuncts of abducibles, we assume that they share the same set of abducibles. Let such set

[2] It is worth noticing, however, that despite this implementation choice, the ALIAS architecture is not strictly bound to the KM proof-procedure: the same high-level features of the system could exploit different abduction algorithms ([Denecker and Schreye, 1998, Kakas, 2000].

[3] In this context, *local* means *performed within the agent that demonstrates the query*.

be

$$\mathcal{A} = \{someone\_else\_killed\_knott, peyer\_killed\_knott\}$$

plus all negated literals. The two agents therefore form a bunch together, $B = \{A_0, A_1\}$.

In order to demonstrate the goal (observation) *:-found_knott's_body*, agent A_0 applies its first clause reducing the goal to *:- peyer_killed_knott*. Then, since *peyer_killed_knott* is abducible the consistency for *peyer_killed_knott* has to be checked within the bunch $\{A_0, A_1\}$. In particular, agent A_1 tries to prove the failure of the integrity constraint $\leftarrow$ *peyer_killed_knott, peyer_has_alibi*. To do this, the failure of *peyer_has_alibi* must be proved: this derivation fails, since A_1's ALP contains the fact *peyer_has_alibi*.

Then, A_0 applies its second clause reducing the goal to *:- someone_else_killed_knott*; again, the consistency check would require to prove the failure of integrity constraints: in this case, only the constraint in A_0 contains the predicate *someone_else_killed_knott* representing an hypothesis against Peyer's guilt. In order to prove its failure, we abduce also the hypothesis *not peyer_killed_knott* so the derivation terminates with success, producing the set of hypothesis

$$\delta = \{someone\_else\_killed\_knott,\ \mathit{not}\ peyer\_killed\_knott\}$$

In the previous example we have considered a pre-defined bunch of two agents for explaining goals with a set of abducible hypotheses which is globally consistent with the knowledge bases of all agents in the bunch. However, in ALIAS more flexible forms for agents composition and coordination are possible, thanks to a *Language for AbductIve Logic Agents* (*LAILA*). In the following we describe it briefly, especially focusing on the LAILA operators used for agent grouping, interaction and coordination.

2.4 The LAILA Language

The *ALIAS* agent behaviour can be expressed by means of the *Language for AbductIve Logic Agents*, LAILA. This language allows to model agent actions and interactions in a logic programming style.

In particular, using the LAILA language, the programmer can express the features of an abductive multi-agent application in a declarative way. In this paper we will focus on agent social behaviour, and on how each agent can request demonstration of goals to other agents in the system. A detailed description of the language can be found in [Ciampolini et al., 2002].

With regard to agent coordination, we introduce two composition operators: the collaborative AND operator (`&`) and the competitive operator (`;`) which can be used by each agent to express and coordinate abductive queries to other (set of) agents. The language provides also a communication operator (>) which is used to submit queries to other agents.

Competitive queries. The competitive operator (`;`) could be exploited to formulate a *disjunction* of alternative queries involving two (or more) abductive agents.

Let us consider the following LAILA competitive query q, formulated by agent A_0:

$$?\ \mathtt{A1} > \mathtt{G}_1\ ;\ \mathtt{A2} > \mathtt{G}_2$$

It means that A_0 must either perform a local abductive derivation for the goal (G_1), *or* ask agent A_1 to demonstrate goal G_2. If several answers are available, the system will select only one of them.

In more detail, the competitive query q causes the following effects:

- A_0 asks A_1 to solve G_1; if G_1 succeeds in A_1, then N ($N > 0$) abductive explanations δ_{1i} (where $i \in [1, \dots N]$), consistent in the bunch $\{A_0, A_1\}$, could be obtained for G_1.

- A_0 asks A_2 to solve G_2; if G_2 succeeds in A_2, then M ($M > 0$) abductive explanations δ_{2j} (where $j \in [1, \dots M]$) consistent in the bunch $\{A_0, A_2\}$, could be obtained for G_2.

- The resulting abductive explanation is either δ_{1i} (for some $i \in [1, \dots, N]$) or δ_{2j} (for some $j \in [1, \dots, M]$).

If both A_1 and A_2 fail, the competitive query fails.

Collaborative queries. The collaborative operator (&) is used to ask queries to different abductive agents with the aim of merging their answers into a unique consistent set of hypotheses. For instance, each query could represent a sub-problem of the problem to be solved.

Let us give an example of collaborative query, issued by agent A_0:

$$?\ \texttt{A1}\ >\ \texttt{G1}\ \&\ \texttt{A2}\ >\ \texttt{G2}$$

It expresses that agent A_0 asks agent A_1 to prove G_1 *and* at the same time A_2 to prove goal G_2; the query will succeed *if and only if* both agents A_1 and A_2 succeed; in particular, it will cause the following effects:

- A_0 asks A_1 to solve $G1$; if $G1$ succeeds in A_1, N ($N > 0$) abductive explanations δ_{1i} ($i \in [1, \dots N]$), consistent in the bunch $\{A_0, A_1\}$, could be obtained for $G1$.

- A_0 asks A_2 to solve $G2$; if $G2$ succeeds in A_2, M ($M > 0$) abductive explanations δ_{2j} ($j \in [1, \dots M]$) consistent in the bunch $\{A_0, A_2\}$, could be obtained for $G2$.

- The abductive explanation for the collaborative query q is therefore a set of hypotheses:

$$\Delta_q = \{\delta_k \mid (\delta_k = \delta_{1i} \cup \delta_{2j}) \ \wedge \ \delta_k \text{ is consistent in } \{A_0, A_1, A_2\}\}$$

If either $A1 > G1$ or $A2 > G2$ fails, the query q fails.

In [Ciampolini et al., 2002] the interested reader can find in detail the operational semantics of the operators cited above, along with the definition of our notion of *consistency*. The declarative semantics of the notion of consistency of a set of hypotheses δ in a bunch $B = \{A_i\}$ of agents is defined as follows:

$$\forall A_i \in B \quad P_i \cup \delta \not\models \bigcup_{j \in S} IC_j$$

In this notation, S is the set of indexes of agents A_i in B, and each agent A_i contains a logic program P_i and a set of integrity constraints IC_i. This definition ensures the absence of conflicts between δ and any integrity constraints in any of B's agents. LAILA allows to implement such semantics with no need to put the agents' knowledge bases together (all the computation is done locally).

2.5 ALIAS implementation

The current implementation of *ALIAS* relies on Jinni [Tarau, 1999], a lightweight, multi-threaded,[4] logic programming language extended with Linda-like primitives,[5] used as a development platform for building multi-agent applications based on logic agents. Although we are aware that a more direct and customized implementation should be more efficient than the current one, we choose Jinni mainly for its high-level and declarative features that allow us to achieve a rapid prototyping of the *ALIAS* system. The adoption of Jinni allows an effective qualitative leap in terms of code readability, homogeneity, easy maintenance, and expressiveness. In fact, both the reasoning and the coordination issues introduced by *ALIAS* are solved by means of logic clauses, leading to the development of a fully declarative design. Moreover, *ALIAS* is platform-independent, since Jinni is written in Java.

3 Multi-agents and logic-based abduction for legal justification: a case study

In this section, we show a possible application of ALIAS in the legal domain. For this purpose, we borrow an example from the literature: the Peyer case. This case has been first cited in [Thagard, 1989] by Thagard, as a case study for showing hypotheses-enforcement via his connectionist program, ECHO, and further reconsidered by at least another project from a different team, doing related research in abduction and its application. In Section 4 we will give a brief review of some related work. Now we would like to introduce the Peyer case and show how it can be modelled in ALP.

3.1 The Peyer Case

The Peyer case is about a twenty-two-year-old San Diego State University honor student, Cara Knott, found strangled and thrown from a 75-foot bridge beside Interstate 15, near San Diego, California, on December 27, 1986. The trial ended in June 1988 in San Diego Superior Court, with a police agent, Craig Peyer of Poway, found guilty of first-degree murder. Thagard's formalization of the problem, published in 1989 and further used by

[4] Jinni makes use of thread technology, which means that different *threads* of execution can be active within the same process. This is important in the case of agents, if we implement each agent with a process, and we need agents that are able to perform multiple actions at the same time, e.g., reasoning to formulate questions to other agents, and on the other hand reasoning to produce answers to questions coming from other agents.

[5] Linda [Gelernter, 1985] is a coordination language used to synchronize the execution of several concurrent processes. Linda provides an abstraction, called *blackboard*, where the processes can post/fetch messages through primitives such as `out` (to post a message) and `in` (to fetch a message). Process synchronization is achieved through a blocking mechanism, that forces a process to suspend its execution after an instruction containing the performative `in`, when waiting for a message that has not been posted yet by another process. Agent coordination in ALIAS is achieved by means of such primitives.

Evidence

[E1]: *Knott's body and car were found on a frontage road near I-15.*
[E2]: *22 young women reported being talked to at length by Peyer after being stopped near where Knott's body was found.*
[E3]: *Calderwood said that he saw a patrol car pull over a Volkswagon like Knott's near I-15.*
[E7]: *Ogilvie said Peyer quizzed her about the case and acted strangely.*
[E8]: *Dotson said Olgivie is a liar.*
[E12]: *Anderson says she saw Peyer wipe off his nightstick in his trunk.*
[E14]: *Bloodstains found on Knott's clothes matched Peyer's blood.*

Figure 1: Excerpt A of Thagard's modeling of the Peyer's case: evidence of Peyer's case.

[Fox and Josephson, 1990], is based on evidence and hypotheses as described, the day following the murder, in the *San Diego Union* and the *San Diego Tribune*. Based on Thagard's report, we will show here how the case can be put in the form of abductive logic programs, representing the knowledge of a multitude of actors (the agents) that actually plaid a role in the Peyer trial, as witnesses, the defendant suspected of committing the murder (Peyer himself) – these being characters from real life – along with an imaginary detective, and especially, an imagined jury whose tilting one way or the other is modelled.

Following on Thagard's formalization, we introduce a number of predicates divided into *evidence predicates*, including testimonies, evidence, findings, matters of fact, etc., *hypotheses*, both for Peyer's guilt and innocence, and *rules* that explain evidence predicates in terms of hypotheses. Being a case study, the formalization of the problem is indeed simplified. We are far away from judicial decision-making may machine; almost nothing could be more controversial. Yet, it is not unconceivable there could be, even outside the ivory towers of academia, real application scenarios, where agents representing juries could for instance abductively reason about the chances, for prosecution, to prove its case: the user would be a prosecutor having the discretion to prosecute or propose a plea bargain. In fact, it would be prohibitive, in terms of resource allocation, to prosecute all cases. Many cases are just dropped. Other suspecs are offered a compromise.

Let us start by giving the formalization of the Peyer's case as in [Thagard, 1989]. In Figure 1 we list a collection of evidence, in Figure 2 the hypotheses and in Figure 3 the case for Peyer's innocence and guilt, expressed in terms of rules that bind evidence with possible explanations. In these figures we adopt the same notation as Thagard [Thagard, 1989] for evidence and hypotheses. In Figure 4 we show the "contradictions", that we express in our framework in terms of integrity constraints.

3.2 Modeling the Peyer case in ALIAS

In this section we model the Peyer case using abductive logic-based agents and the ALIAS multi-agent architecture.

As already anticipated, the *dramatis personæ*, i.e., the characters that are relevant to solve the case, are modelled by means of agents. We refer to some standard roles that an agent can play, for instance, a detective, a witness, a suspect, a member of a jury, etc. In the examples that we show in the sequel, we will call d the detective, w_X the witnesses, and s_X a suspect, where X is the name of the particular witness or suspect. In this simplified example we do

Hypotheses that Peyer is guilty	**Hypotheses that Peyer is innocent**
[G1]: *Peyers killed Knott.*	**[I1]:** *Someone other than Peyer killed Knott.*
[G4]: *Peyer pulled Knott over.*	**[I4]:** *Ogilvie lied.*
[G6]: *Peyer like to pull over young women.*	**[I4A]:** *Ogilvie is a liar.*
[G7]: *Peyer had a bloody nightstick.*	**[I7]:** *Anderson was mistaken about the nightstick.*

Figure 2: Excerpt B of Thagard's modeling of the Peyer's case: list of possible explanations of evidence.

The case for Peyer's guilt	**The case for Peyer's innocence**
G7 $\leftarrow$ **G1**	**E1** $\leftarrow$ **I1**
E1 $\leftarrow$ **G1**	**E7** $\leftarrow$ **I4**
E2 $\leftarrow$ **G6**	**E8** $\leftarrow$ **I4A**
E3 $\leftarrow$ **G4**	**I4** $\leftarrow$ **I4A**
E7 $\leftarrow$ **G1**	**E12** $\leftarrow$ **I7**
E12 $\leftarrow$ **G7**	
E14 $\leftarrow$ **G1**	

Figure 3: Excerpt C of Thagard's modeling of the Peyer's case: relationships between evidence and possible explanations.

Contradictions

$\leftarrow$ **G1, I1**
$\leftarrow$ **G7, I7**

Figure 4: Excerpt D of Thagard's modeling of the Peyer's case: contradictions between hypotheses.

not explicitly model the members of a jury, but we rather focus on a pre-trial phase, and on the effort made by factfinders in combining testimonies and establishing relationships among witnesses, pieces of evidence and testimonies.

In our framework, we model the *data* that are relevant to make any kind of inference as logical terms. We distinguish between a *piece of evidence* and a *testimony*. A piece of evidence is, for instance, the one identified with **E14** in [Thagard, 1989] (v. Figure 1): "*Bloodstains found on Knott's clothes matched Peyer's blood*". Pieces of evidence are mapped as non-abducible predicates, being incontrovertible facts. They can be modeled either as facts, that are assumed to be true, or as observations, that must be explained y means of suitable hypotheses.

For instance, "*22 young women reported being talked to at length by Peyer*"[6] (**E2**) is a fact, that can be used to explain more facts. "*Knott's body and car were found on a frontage road near I-15*" (**E1**), instead, is an observation that can be explained by means of hypotheses, like: "*Peyers killed Knott*" (**G1**), or "*Someone else killed Knott*" (**I1**).

Testimonies do not represent facts. Witnesses can be more or less reliable. What is true in a testimony is indeed the act of uttering it, but not necessarily the content of the testimony. If Calderwood declares "*I saw a patrol car pull over a Volkswagen like Knott's near I-15*", this does not necessarily imply that a patrol car pulled over any Volkswagen ever, while it does indeed imply that Calderwood (`w_calderwood`) released such a testimony. In our representation of a case, we will therefore distinguish between a testimony and content of a testimony, and consider the content of a testimony as a defeasible hypothesis that can be assumed true, or false.[7] We will map testimonies into atoms starting by "`t`", and hypotheses as atoms starting by "`h`": for instance, "*Calderwood said that a patrol car pulled over a Volkswagen like Knott's near I-15*" will be mapped into the atom `t_patrol_car_pulled_over_vw`, while the hypothesis "*a patrol car pulled over a Volkswagen like Knott's near I-15*" will be mapped into `h_patrol_car_pulled_over_vw`.

Operationally, as we will see in the sequel, this kind of hypotheses which can emerge from direct testimonies will be generated during the multi-agent abductive inference process by the witness agents themselves. For instance, a witness will reply to a question about some aspects of the case, by means of rules of the kind:

$$\texttt{ask(Aspect)} \leftarrow \texttt{h\_X}$$

For the example above, the agent acting on the behalf of Calderwood will have the following rule, which expresses that agent's ability to answer a question about the car:

$$\texttt{ask(car)} \leftarrow \texttt{h\_patrol\_car\_pulled\_over\_vw}$$

In [Thagard, 1989], a rule like this is represented by means of a proposition (see Figure 2), and there is no distinction between the content of a proposition and the act of uttering the proposition itself.

[6] According to several testimonies, Peyer seemed to take over young women and talk to them, seemingly exploiting the rights given by his status of police agent for personal purposes.

[7] In the judicial context, there are in a witness's testimony parts that – if proven to be false – completely invalidate it, and there are parts that – if they are false – do not completely invalidate it because they are irrelevant, especially if they do not show an intention to lie *tout-cours* ("Me? I am a good guy. I come from a noble old family". This kind of fibs is not fatally invalidating to the entire testimony of the given witness).

The objective of a detective d in this setting is that of providing objective explanations of observations, thus finding for instance the responsible persons of a crime. This is done in our multi-agent framework in the following way: The detective triggers the explanation of testimonies by means of constraints of the kind:

$$\leftarrow \mathtt{h\_t\_X}, \mathtt{not}\ \mathtt{expl}(\mathtt{h\_t\_X})$$

where `expl` is a meta-predicate on abducibles, defined by rules that represent the detective's reasoning activity.

For instance:

```
← h_patrol_car_pulled_over_vw,
  not expl(h_patrol_car_pulled_over_vw).
expl(h_patrol_car_pulled_over_vw) ←
      h_peyer_pulled_over_knott.
expl(h_patrol_car_pulled_over_vw) ←
      h_peyer_pulled_over_another_vw.
expl(h_patrol_car_pulled_over_vw) ←
      h_another_patrol_car_pulled_over_knott.
```

and so forth.

Modelling reliability

This schema can be also adopted in the case of witnesses that predicate about other witnesses' reliability. The Peyer case provides us once more with an example of this: a witness, Dotson, declares that another witness, Ogilvie, is a liar. Inside Dotson's knowledge base we find the rule:

$$\mathtt{ask}(\mathtt{ogilvie}) \leftarrow \mathtt{h\_ogilvie\_liar}.$$

and inside the detective, the constraint that puts such hypothesis into relationship with its explanation.

```
← h_ogilvie_liar, not expl(h_ogilvie_liar).
expl(h_ogilvie_liar) ← not h_rel_ogilvie.
```

We imagine the solution of a case such as Peyer's as a dynamic process, in which we can, for instance, add more questions to witnesses, relevant facts, or evidence depending on the current state of the trial and on our knowledge.

Once we obtain a certain conjunct of hypotheses Δ, with a particular group of agents, we can check it against an enlarged group of agents including more witnesses, for example in order to verify their reliability. Here comes in hand the disjunctive LAILA operator (competitive operator):

$$\Delta\ \ \&\ \ (\mathtt{w\_X}\ >\ \mathtt{ask}(\mathtt{Y})\ ;\ \mathtt{not}\ \mathtt{h\_rel\_Y}).$$

In the case of Ogilvie/Dotson, if I am "certain" that Ogilvie is a reliable witness, I can query Dotson about Ogilvie herself in the following way:

$$\{\texttt{h\_rel\_ogilvie}\} \ \texttt{\&} \ (\texttt{w\_dotson} > \texttt{ask}(\texttt{ogilvie})\ ;\ \texttt{not}\ \texttt{h\_rel\_dotson}).$$

This LAILA query would produce the following effects: if we follow the first branch of the competitive query we would fail, since from

$$\texttt{w\_dotson} > \texttt{ask}(\texttt{ogilvie})$$

we derive the hypothesis `h_ogilvie_liar` from Dotson's query, that would activate the detective's integrity constraint which forces him to explain

$$\texttt{h\_ogilvie\_liar:expl}(\texttt{h\_ogilvie\_liar})$$

. This together with the other integrity constraint, would generate the hypothesis `h_rel_ogilvie` which clashes with Δ.

The second branch, instead, succeeds, leading to a specialization of the initial hypothesis:

$$\Delta' = \{\texttt{h\_rel\_ogilvie}, \texttt{not}\ \texttt{h\_rel\_dotson}\}$$

. I can therefore conclude that Dotson is not a reliable witness.

In this way, the detective is able to collect all the hypotheses that lead to a possible reconstruction of the events, and he will try to prove it in collaboration with the group of witnesses that proved so far reliable.

For instance, a possible logical justification of the case for Peyer's guilt is:

```
expl(h_peyer_killed_knott) ←
    h_peyer_pulled_over_knott,
    h_peyer_hit_knott_with_nightstick.
```

This justification is supported by specific testimonies (Calderwood's about the car – **E3** – Anderson's about the nightstick – **E12** – and so on). Therefore we can refine the detective's knowledge base by adding the following query:

```
d > expl(h_peyer_killed_knott) &
d > reliable_witnesses &
w_calderwood > ask(car) &
w_anderson > ask(nightstick).
```

where the definition of `reliable_witnesses` is updated depending on the current beliefs of the detective about the witness reliability. For instance:

```
reliable_witnesses ←
    h_rel_calderwood,
    h_rel_anderson,
    not h_rel_dotson.
```

Inconsistencies and collusion

It is possible in this way also to prove the inconsistencies of various testimonies: for instance, Peyer says that he did not have a bloody nightstick, while Anderson says that she saw Peyer wiping off the nightstick on his trunk:[8]

```
% Peyer's knowledge base
ask(nightstick) ←
   h_not_bloody_nightstick.

% Anderson's knowledge base
ask(nightstick) ←
   h_wipe_nightstick.

% detective's knowledge base
← h_wipe_nightstick, not expl(h_wipe_nightstick)
← h_not_bloody_nightstick, h_bloody_nightstick
expl(h_wipe_nightstick) ← h_bloody_nightstick
```

If the detective asks:

```
s_peyer > ask(nightstick) & w_anderson > ask(nightstick)
```

it obtains the following:

```
s_peyer > ask(nightstick)
```

produces `h_not_bloody_nightstick`, which by the detective's integrity constraints generates `not h_bloody_nightstick`. On the right hand side,

```
w_anderson > ask(nightstick)
```

produces the hypothesis `h_wipe_nightstick`, that by the detective's integrity constraints generates `expl(h_wipe_nightstick)` and therefore `h_bloody_nightstick`. This hypothesis is in contrast with the previous one obtained from Peyer. If both are the only open derivation branches in this query, it indeed fails, which can be interpreted as an inconsistency in the two persons' testimonies, provided that they are properly modelled in the system.

We see how the inquiry or trial, as it progresses in its various phases, can highlight inconsistencies among the testimonies. If `h_rel_anderson` is in the set of witnesses that I currently consider reliable (defined by `reliable_witnesses`), I can obtain from a query such as that above specified a new Δ that gives reason to Anderson, and I can use the new hypothesis `not h_rel_peyer` to update the definition of `reliable_witnesses`. Different definitions of `reliable_witnesses` can lead to different conclusions.

By proceeding in a similar way to that used to find out inconsistencies, we can use LAILA to determine groups of agents that agree between themselves, but that are not consistent with some evidence that the detective knows. This can be used, case by case, to raise some hypotheses that assume what the agreements between testimonies or suspects are.

[8] In the sequel, comments to pieces of code are preceded by the symbol % .

3.3 Some guidelines about modeling case using ALIAS

To sum up, we give some guidelines about how to model a case in ALIAS.

- *Agents.* Used to model the *dramatis personæ* of a case. Each agent incorporates an abductive logic program for internal reasoning, plus a LAILA program for interactions and queries that involve more that one agent. An abductive logic program contains abducibles, definitions, and integrity constraints.

- *Abducibles.* Used to model the content of testimonies and hypotheses, the latter being either reliability hypotheses or other hypotheses adding during the solution of the case (e.g. hypotheses of agreement among other agents)

- *Definitions.* In the detective's program, used to explain hypotheses and to define a set of possible hypotheses that can be used as initial assumptions for a given query. For instance, an initial Δ, or a set of reliable agents. In a witness's program they define the `ask` predicate, used to answer to queries about specific aspects of the case. There could be indeed more definitions that represent the internal reasoning of individuals.

- *Integrity constraints.* In addition to implicit constraints used by the abductive proof, e.g.,

  $$\leftarrow \texttt{h\_bloody\_nightstick}, \texttt{not}\ \texttt{h\_bloody\_nightstick},$$

  integrity constraints can be used to enforce the explanation of given hypotheses, e.g.

  $$\leftarrow \texttt{h\_bloody\_nightstick}, \texttt{not}\ \texttt{expl}(\texttt{h\_bloody\_nightstick}).$$

  As in [Thagard, 1989], integrity constraints are used also to prevent from assuming contradictory hypotheses, e.g.,

  $$\leftarrow \texttt{h\_peyer\_killed\_knott}, \texttt{not}\ \texttt{expl}(\texttt{h\_someone\_else\_killed\_knott}).^9$$

  There could be indeed more integrity constraints in the body of other agents, and not only in the detective's body. For instance, a witness who declares that two hypotheses cannot be assumed to be true at the same time (e.g., Peyer declares that he was not in the same place as Knott: $\leftarrow \texttt{h\_peyer\_on\_I15}, \texttt{h\_knott\_on\_I15}$).

4 Related Work

Our work relates to several research areas of both artificial intelligence and law, and, of course, of a discipline almost as old (some would say: older[10]) than AI, i.e., "AI & Law" In this section, we mainly focus on the work done in artificial intelligence, with particular regard to abduction, logic programming and multi-agent systems.

Among the other approaches to modelling jurisdictional cases by means of abductive frameworks, Thagard in [Thagard, 1989] describes a framework (ECHO) for explaining sets of propositions that he calls "evidence". Pieces of evidence are connected together by means of explanatory/contradictory relationships. We already discussed along the paper the main

[10]This is because deontic logic has played an important role in AI & Law since 1970s (e.g. see [Nissan, 1998]), so some would trace it back to legal philosopher's introduction of deontic logic in 1940s.

differences between Thagard's modeling of Peyer's case and ours. In particular, we saw how ALIAS/LAILA is based on an abductive logic programming-based framework where the knowledge is distributed, and allows for asking specific queries to the agents/reasoners, without having to consider all the facts in the model as a whole. This allows us to establish specific relations between pieces of evidence and testimonies, and groups of agents. ECHO, on the other hand, is a monolithic system, which does not have the concept of *dramatis personæ*, and is based on a connectionist model which adopts a continuous truth value. The evaluation of the model leads to a unique answer that shows which hypotheses are explanatory for the initial evidence, and which are not, based on a certain fixed threshold.

Josephson et al. also developed an abductive inference system called *Peirce* [Fox and Josephson, 1990, Punch III et al., 1990], a domain independent abductive reasoner that also makes use of a connectionist model. Peirce generates "highly confidencial" hypotheses starting from an input constituted by a set of findings to be explained, together with some initial scores. As opposed to Thagard's ECHO, Peirce allows for assigning initial likelihood values to the hypotheses, and is based on the concept of "essentials", i.e., hypotheses that must be assumed true, since they are the only ones that explain some specific findings. Peirce's algorithm progressively enlarges the set of essentials and rules out hypotheses that are incompatible with them. The authors claim that this is faster than ECHO.

Independently of Thagard's approach, and about the same time in the late 1980s, the ALIBI project was developed by a team led by Nissan. Alternative, exonerating explanations for charged behavior (accusations) are proposed by an automated planner impersonating the accused person; see [Kuflik et al., 1991, Fakher-Eldeen et al., 1993, Nissan, 1996, Nissan and Rousseau, 1997, Nissan, 2000]. When ALIBI tries to explain out damning evidence, the concocted explanation may be funny. The COLUMBUS model is on funny explanations in humoristic literary text [Nissan, in press].

In our work, we focus on the expressive power of the agent metaphor. This allows to propose a model where different pieces of software have different roles, which reflect their role in the real world case. This has several advantages, not only in terms of model, but also in terms of flexibility and compositionality of the whole framework. We could imagine agents as components that can be re-used to deal with different cases: for instance, agents enclosing whole *corpora* of laws/regulations of some countries, or agents that represent organizations, or that reproduce the behaviour of other pieces of software. Although our implementation is only at a prototypical stage, we believe that the multi-agent approach based on abductive logic programming has the potential to become a flexible tool to model real world cases, provided a "safe" niche is found, i.e., not for involvement in reaching verdicts in court, but in either theoretical analyses of real cases of judicial decision-making [Hastie, 1993, Kadane and Schum, 1996], or for the analysis of costs and benefits of whether prosecution should take a case to court at all. Quantitative models of judicial decision-making are not necessarily probabilistic [Shimony and Nissan, 2001], yet they most often have been, and this in turn is the object of a heated controversy [Nissan, 2001a, Nissan, 2001b, Allen and Redmayne, 1997].

With regard to abductive logic programming, this approach has already been adapted to legal reasoning [Satoh, 1996, Satoh, 1998], which we consider a further confirmation with respect to choices we made. In particular, we would like to cite the work by Satoh in [Satoh, 1996] where the author proposes a method to translate case-based reasoning into abductive logic programming and shows how it is possible to model itentions of people involved in a trial to explain given observations. In [Satoh, 1998], Satoh implements a dynamic notion of similarity of cases in legal reasoning. The system translates a legal case into an abductive logic program, possibly retrieves a similar case, and produces an explanation why the current case is similar to the retrieves one. This work could be used a support for making a library of cases, that could help in modelling legal cases and use them in ALIAS. Argumentation theory has also been estensively used in legal reasoning for instance to reconcile conflicting arguments in a similar way to what we do by means of logic programming [Kowalski and Toni, 1994]. We are not going to survey work on argumentation since it is outside the scope of this paper; However, the interested reader in the relationship between abduction and argumentation can refer to [Kakas et al., 1998]. For recent surveys on abduction in Logic Programming, see [Kakas and Denecker, 2002] and [Kakas et al., 1998]. For a discussion on other work related to ALIAS/LAILA, see [Ciampolini et al., in press, Ciampolini et al., 2002].

5 Conclusion

In this paper we have presented a new multi-agent approach for legal justification. We have shown that the process of explaining pieces of evidence and other observation in a trial could profit from the use of logic-based abduction within the framework of a multi-agent architecture. In summary the main features of this approach are:

- the declarative representation of trial knowledge, by means of abductive logic programming, allows to express knowledge and reasoning criteria at a very high level.
- mapping distinct characters of a trial into distinct agents allows the modular and dynamic composition of distinct sources of knowledge; in this way it is possible to perform collaborative/competitive reasoning on trial data, using the knowledge of two or more *dramatis personæ* to obtain a consistent set of hypotheses that explains the given data;
- agents collaborative explanation allows also the detection of incoherence and collusion among trial characters, and can produce reliability hypotheses about people implicated in the trial; these reliability hypotheses could drive detective's enquiry activity towards the final solution of the case.

As a matter of future work, we plan to extend the ALIAS architecture towards the introduction of probability values as scores associated to collaborative/competitive abductive explanations. In this respect, we have already studied the integration of the Probabilistic Horn Abduction theory [Poole, 1993] with abductive logic agents, in particular with the aim of performing distributed medical diagnosis. Therefore we will adapt that approach, if possible, to the case of legal justification and then we will possibly implement it by extending the ALIAS implementation. We are aware, however, of Bayesianism being highly controversial in legal evidence theory among legal scholars: as pointed out by a leading "Bayesio-skeptic", [Allen, 2001], the real challenge for the application of AI to legal evidence is in developing such models of plausiblity that would step forward, leaving behind the controversial grounds of probabilistic reasoning.

Acknowledgements

This work was supported by the SOCS project, funded by the CEC, contract IST-2001-32530. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

[Allen, 2001] Allen, R. (2001). Artificial intelligence and the evidentiary process: the challenges of formalism and computation. *Artificial Intelligence and Law*, 9(2/3):99–114.

[Allen and Redmayne, 1997] Allen, R. and Redmayne, M., editors (1997). *Bayesianism and Juridical Proof*, volume 1 of *Special issue of The International Journal of Evidence and Proof*.

[Brogi et al., 1997] Brogi, A., Lamma, E., Mancarella, P., and Mello, P. (1997). A unifying view for logic programming with non-monotonic reasoning. *Theoretical Computer Science*, 184:1–49.

[Ciampolini et al., in press] Ciampolini, A., Lamma, E., Mello, P., Toni, F., and Torroni, P. (in press). Co-operation and competition in *ALIAS*: a logic framework for agents that negotiate. *Annals of Mathematics and Artificial Intelligence, special issue in Computational Logic and Multi-Agency*.

[Ciampolini et al., 2002] Ciampolini, A., Lamma, E., Mello, P., and Torroni, P. (2002). LAILA: A language for coordinating abductive reasoning among logic agents. *Computer Languages*, 27(4):137–161.

[Clark, 1978] Clark, K. L. (1978). *Negation as Failure*, chapter Logic and Databases. Plenum Press, New York.

[Console et al., 1991] Console, L., Dupré, D. T., and Torasso, P. (1991). On the relationship between abduction and deduction. *Journal of Logic and Computation*, 1(5):661–690.

[Cox and Pietrzykowski, 1986] Cox, P. T. and Pietrzykowski, T. (1986). Causes for events: Their computation and applications. In *Proceedings CADE-86*, page 608.

[Denecker and Schreye, 1998] Denecker, M. and Schreye, D. D. (1998). SLDNFA: an abductive procedure for abductive logic programs. *Journal of Logic Programming*, 34(2):111–167.

[Dung, 1995] Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357.

[Eshgi and Kowalski, 1989] Eshgi, K. and Kowalski, R. A. (1989). Abduction compared with negation by failure. In Levi, G. and Martelli, M., editors, *Proceedings 6th International Conference on Logic Programming*, page 234. MIT Press.

[Fakher-Eldeen et al., 1993] Fakher-Eldeen, F., Kuflik, T., Nissan, E., Puni, G., Salfati, R., Shaul, Y., and Spanioli, A. (1993). Interpretation of imputed behavior in ALIBI (1 to 3) and SKILL. *Informatica e Diritto*, 2(1/2):213–242.

[Fox and Josephson, 1990] Fox, R. and Josephson, J. (1990). Peirce-IGTT: A domain-independent problem solver for abdcutive assembly. Technical report, The Laboratory for Artificial Intelligence Research, The Ohio State University.

[Fung and Kowalski, 1997] Fung, T. H. and Kowalski, R. A. (1997). The IFF proof procedure for abductive logic programming. *Journal of Logic Programming*.

[Gelernter, 1985] Gelernter, D. (1985). Generative communication in linda. *ACM Toplas*, 7(1):80–112.

[Gini and Ishida, 2002] Gini, M. and Ishida, T., editors (2002). *Proceedings of the first Joint Conference on Autonomous Agents and Multiagent Systems, Bologna, Italy*. ACM Press.

[Hastie, 1993] Hastie, R. (1993). *Inside the Juror: The Psychology of Juror Decision Making*. Cambridge University Press, Cambridge, UK.

[Kadane and Schum, 1996] Kadane, J. B. and Schum, D. A. (1996). *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. Wiley, Chichester, UK.

[Kakas and Denecker, 2002] Kakas, A. and Denecker, M. (2002). Abduction in logic programming. In Kakas, A. and Sadri, F., editors, *Computational Logic: Logic Programming and Beyond. Part I*, number 2407 in LNAI, pages 402–436. Springer Verlag.

[Kakas, 2000] Kakas, A. C. (2000). ACLP: integrating abduction and constraint solving. In *Proceedings NMR'00, Breckenridge, CO*.

[Kakas et al., 1998] Kakas, A. C., Kowalski, R. A., and Toni, F. (1998). The role of abduction in logic programming. *Handbook of Logic in AI and Logic Programming*, 5:235–324.

[Kakas and Mancarella, 1990a] Kakas, A. C. and Mancarella, P. (1990a). Generalized stable models: a semantics for abduction. In *Proceedings 9th European Conference on Artificial Intelligence*. Pitman Pub.

[Kakas and Mancarella, 1990b] Kakas, A. C. and Mancarella, P. (1990b). On the relation between Truth Maintenance and Abduction. In Fukumura, T., editor, *Proceedings of the first Pacific Rim International Conference on Artificial Intelligence, PRICAI-90, Nagoya, Japan*, pages 438–443.

[Kowalski, 1990] Kowalski, R. A. (1990). Problems and promises of computational logic. In *Proceedings Symposium on Computational Logic*, pages 1–36. Springer-Verlag.

[Kowalski and Toni, 1994] Kowalski, R. A. and Toni, F. (1994). Argument and reconciliation. In *Proceedings of Workshop on Legal Reasoning International Symposium on FGCS Tokyo*.

[Kuflik et al., 1991] Kuflik, T., Nissan, E., and Puni, G. (1991). Finding excuses with ALIBI: alternative plans that are deontically more defensible. *Computers and Artificial Intelligence*, 10(4):297–325.

[Lloyd, 1987] Lloyd, J. W. (1987). *Foundations of Logic Programming*. Springer-Verlag.

[Nissan, 1996] Nissan, E. (1996). From ALIBI to COLUMBS: the long march to self-aware computational models of humor. In *Proceedings of the 12nd Twente Workshop on Language Technology*, pages 69–85.

[Nissan, 1998] Nissan, E. (1998). Advances in deontic logic. *Computers and Artificial Intelligence*, 14(4):392–400.

[Nissan, 2000] Nissan, E. (2000). Artificial intelligence and criminal evidence: a few topics. In *Proceedings of the 2nd International Conference on New Trends in Criminal Investigation and Evidence*, pages 495–521. Intersentia, Antwerpen.

[Nissan, 2001a] Nissan, E. (2001a). The bayesianism debate in legal scholarship. *Artificial Intelligence and Law*, 9(2/3):199–214.

[Nissan, 2001b] Nissan, E. (2001b). Can you measure circumstantial evidence? the background of probative formalisms for law. *Information and Communication Technology Law*, 10(2):231–245.

[Nissan, in press] Nissan, E. (in press). The COLUMBUS model (parts I & II). *The International Journal of Computing Anticipatory Systems*.

[Nissan and Rousseau, 1997] Nissan, E. and Rousseau, D. (1997). Towards ai formalisms for legal evidence. In Raś, Z. and Skowron, A., editors, *Foundation of Intelligent Systems: Proceedings of the 10nd Symposium, ISMIS'97*, pages 328–337. Springer-Verlag.

[Poole, 1988] Poole, D. L. (1988). A logical framework for default reasoning. *Artificial Intelligence*, 36(1):27–47.

[Poole, 1993] Poole, D. L. (1993). Logic programming, abduction and probability:a top down anytyme algorithm for estimating prior and posterior probabilities. *Artificial Intelligence*, 64:81–129.

[Punch III et al., 1990] Punch III, W., Tanner, M., Josephson, J., and Smith, J. (1990). Peirce: A tool for experimenting with abduction. *IEEE Expert*, 5(5):34–44.

[Satoh, 1996] Satoh, K. (1996). Translating case-based reasoning into abductive logic programming. In *Proceedings European Conference on Artificial Intelligence (ECAI)*. John Wiley & Sons, Ltd.

[Satoh, 1998] Satoh, K. (1998). Using two level abduction to decide similarity of cases. In *Proceedings 13th European Conference on Artificial Intelligence*, pages 398–402. John Wiley and Sons, Chichester.

[Shimony and Nissan, 2001] Shimony, S. E. and Nissan, E. (2001). Kappa calculus and evidential strength: a note on åqvist's logical theory of legal evidence. *Artificial Intelligence and Law*, 9(2/3):153–163.

[Tarau, 1999] Tarau, P. (1999). Jinni: Intelligent mobile agent programming at the intersection of java and prolog. In *Proceedings PAAM'99, London, UK*.

[Thagard, 1989] Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, (3):435–502.

[Toni, 1995] Toni, F. (1995). A semantics for the Kakas-Mancarella procedure for abductive logic programming. In Alpuente, M. and Sessa, M. I., editors, *Proc. GULP'95*, pages 231–242.

[van Emden and Kowalski, 1976] van Emden, M. H. and Kowalski, R. A. (1976). The semantics of predicate logic as a programming language. *Journal of the ACM*, 23(4):733–742.