Investigating Stories in a Formal Dialogue Game

Floris BEX [a], Henry PRAKKEN [a, b]

[a] *Centre for Law & ICT, Faculty of Law, University of Groningen*
[b] *Department of Information and Computing Sciences, Utrecht University*

Abstract In this paper we propose a formal dialogue game in which two players aim to determine the best explanation for a set of observations. By assuming an adversarial setting, we force the players to advance and improve their own explanations as well as criticize their opponent's explanations, thus hopefully preventing the well-known problem of 'tunnel vision'. A main novelty of our approach is that the game supports the combination of argumentation with abductive inference to the best explanation.

1. Introduction

In the literature, two main approaches to reasoning about factual issues in criminal cases are the story-based and the argument-based approach. Some authors [1] argue that legal reasoning about evidence is mainly done by constructing and analysing evidential arguments from the sources evidence to the events that are to be proven. Other authors [13] argue that legal reasoning with evidence is instead done by constructing different stories around the evidence and then analysing and comparing these stories. In previous work ([5], [6]), we have shown that these approaches can be combined in one formal framework, namely as a combination of the AI formalisms of abductive inference to the best explanation (IBE) and defeasible argumentation. Stories about what might have happened in a case are represented as hypothetical explanations and links between stories and the available evidence are expressed with evidential generalizations that express how parts of the explanations can be inferred from evidential sources with defeasible argumentation. Combining the argumentative and IBE approach in this way solves some of the problems of the separate approaches

A limitation of our previous work is that it discusses only a static viewpoint: a framework is provided for the current status of an argumentative and story-based analysis of a case, and the dynamics of developing and refining such an analysis are not discussed. In this paper, we aim to model these dynamics in a formal dialogue game. In this dialogue game, it should be possible to build, critically analyse and change stories and their supporting arguments.

The game we propose is for dialogues in which crime analysts aim to determine the best explanation for a set of observations. Despite this cooperative goal of the dialogue participants, we still assume an adversarial setting in which the protocol is designed to motivate the players to 'win'. Thus we hope to prevent the well-known problem of 'tunnel vision' or confirmation bias, by forcing the participants to look at all sides of a case. A main novelty of our approach, motivated by our previous work in

[5] on the static aspects of crime investigation, is that the game supports the combination of argumentation with IBE.

The rest of this paper is organized as follows. In section 2, we summarize our combined framework as developed in [5]. In section 3 we present our formal dialogue game and in section 4 we apply it to a simple example. Finally, in section 5 we conclude with a discussion and some ideas for future research.

2. Argumentative Story-based Analysis of Reasoning with Evidence

In this section, the framework for argumentative story-based analysis of reasoning with evidence as proposed in [5] will be summarized. This formal framework combines a logic for defeasible argumentation with a logical model of abductive inference to the best explanation (for an overview of abductive reasoning see [10]). We first discuss our combined logical framework, followed by a short example. Then we will argue that this combined approach solves some of the problems of purely argumentative and purely IBE approaches to evidential reasoning.

The basic idea of the combined approach is as follows. A logical model of abductive IBE takes as input a causal theory and a set of propositions that has to be explained, the *explananda*, and produces as output a set of hypotheses that explain the explananda in terms of the causal theory. The combination of hypotheses and causal theory can be seen as a story about what might have happened. These hypothetical stories can then be compared according to the extent to which they conform to the evidence in a case. This evidence is connected to the stories by defeasible arguments from evidential sources (e.g. witness testimonies). Defeasible arguments are also used to reason about the plausibility of a story: the causal rules of the causal theory are not just given but their applicability can become the subject of an argumentation process. This definition of stories as causal networks is not entirely new: Pennington and Hastie [13] also defined stories as causal networks, following earlier influential research by Schank and Abelson [18]. Note that we use a naïve interpretation of causality; sometimes a causal link does not represent a much stronger relation than temporal precedence. This allows us to model a story as a simple, chronologically ordered sequence of events.

A framework for evidential reasoning $ER = (C, A)$ is a combination of a causal-abductive framework C and an evidential argumentation framework A. The underlying logic L of this framework consists of the inference rules of classical logic combined with a defeasible modus ponens rule for a conditional operator $\Rightarrow$ for defeasible generalizations. The generalizations used in C and A (see below) are formalized with this connective: $g_i: p_1 \wedge ... \wedge p_n \Rightarrow q$. Here g_i is the name of the generalisation and $p_1 ... p_n$ and q are literals. The type of generalisation is indicated with a subscript: $\Rightarrow_E$ denotes an evidential generalisation and $\Rightarrow_C$ denotes a causal generalisation. For example, Smoke $\Rightarrow_E$ Fire says that smoke is evidence of fire, while Fire $\Rightarrow_C$ Smoke says that fire causes smoke.

The *argumentation framework* is a pair $A = (G, I)$, where G is a set of evidential generalizations and I is a set of input facts, where $I_E \subseteq I$ is the set of sources of evidence in a case. This set of evidence I_E is different from other input facts in that the sources of evidence in I_E cannot be attacked by arguments. The other elements in I are propositions that denote 'general knowledge', opinions or ideas which may be open to

discussion and which are not generalizations of the form 'if…then…' (e.g. 'Hillary Clinton will probably be the next U.S. president' or 'the idea that people from Suriname rob supermarkets more often than Dutch people is based on prejudice').

The logic for this framework is very similar to the logic underlying the ASPIC inference engine [1], which in turn combines Pollock's [14] ideas on a tree structure of arguments and two notions of rebutting and undercutting defeat with Prakken & Sartor's [17] rule language and their argument game for Dung's [8] grounded semantics. The set I and the evidential generalizations from G allow us to build evidential arguments by taking elements from I and the generalizations as premises and chaining applications of defeasible modus ponens into tree-structured arguments. Such an *evidential argument* is a finite sequence of lines of argument, where a line is either a proposition from I, a generalization from G or the result of an application of the defeasible modus ponens to one or more previous lines. $Args(A)$ is the set of all well-formed arguments in A.

An argument can defeat another argument by rebutting or undercutting the other argument. Two arguments *rebut* each other if they have the opposite conclusion. An argument AR_1 *undercuts* another argument AR_2 if there is a line $\neg g_i$ in argument AR_1 and a line in argument AR_2, which is obtained from some previous lines in AR_2 by the application of defeasible modus ponens to g_i.

For a collection of arguments and their binary defeat relations, the dialectical status of the arguments can be determined: arguments can be either *justified*, which means that they are not attacked by other justified arguments that are stronger, or *overruled*, which means that they are attacked by one or more other stronger arguments that are justified, or *defensible*, which means that they are neither justified nor overruled. Note that in the present paper, we will not discuss the relative strength between arguments.

The *abductive framework* is a tuple $C = (H, T, F)$. Here, T is the *causal theory* which contains all the causal generalizations from the different stories. H is a set of *hypotheses*, propositions with which we want to explain the explananda. F is the set of *explananda*, propositions that have to be explained. A set of hypotheses H and a causal theory T can be used to explain propositions:

(**explaining**) $H_i \cup T_i$, where $H_i \subseteq H$ and $T_i \subseteq T$, *explains* a set of propositions E iff
1. $\forall e$: If $e \in E$ then:
    - $H_i \cup T_i \vdash e$; and
    - $H_i \cup T_i$ is consistent.
2. There is no justified argument in $Args(A)$ for the conclusion $\neg g$, where $g \in T_i$.

Here $\vdash$ stands for logical consequence according to the set of all deductive inference rules extended with modus ponens for $\Rightarrow$. Condition (1) of this definition is standard in logical models of abduction but condition (2) is new and makes it possible to attack causal generalizations of dubious quality in the explanation with an argument: $H_i \cup T_i$ does not explain E if one of the generalizations in T is attacked by a justified argument.

The idea of our dialogue game is that during a dialogue the players jointly and incrementally build a framework, which can contain several alternative explanations for the explananda. Moreover, these explanations can be extended during the dialogue, for instance, by giving a further explanation for a hypothesis. Therefore we must be able to identify at each step of the dialogue the explanations for F.

(**explanation**) Given a framework ER, an explanation $S = H_i \cup T_i$ is an explanation for the explananda F iff:
1. S explains F; and

2. $H_i \subseteq H$ contains only initial causes; and
3. T_i is a minimal subset of T.

Initial causes are propositions that are not a conclusion of a causal rule in T. This ensures that an explanation is considered from beginning to end. The condition that T_i is a minimal subset of T ensures that two explanations for F are really seen as two different explanations. The set of all explanations for F in a framework is denoted as *Expl*(*ER*). If there is more than one explanation for the explananda, they must be compared according to their plausibility and their conformity to the evidence in a case.

The plausibility of a story is often judged by looking at the plausibility of its underlying generalizations (cf. [19]). Two kinds of generalizations are important to consider: the causal generalizations in the explanation and the evidential generalizations in the arguments linking the evidence to the story. The plausibility of the causal generalizations is ensured by point 2 of the definition of explaining on the previous page. In the same way, the plausibility of the evidential generalizations in the arguments is ensured by allowing arguments to be attacked and defeated.

As to an explanation's conformity to the evidence in a case, we recognize three criteria. The first of these is *evidential coverage*, which stands for the number of sources of evidence covered by an explanation. The second is *evidential contradiction*, which stands for the number of events in the explanation contradicted by evidential arguments and the third is *evidential support*, which stands for the number of events in a story supported by evidential arguments. Evidential coverage was first mentioned in [13] and the other criteria were mentioned in [6]. Because in this paper the focus is on the dialogue game, we only formally define evidential coverage.

(**evidential coverage**) The *evidential coverage* of an explanation S, denoted as ec_S, is the total number of sources of evidence that are *covered* by an explanation, where:

- a source of evidence $p \in I_E$ is covered by an explanation S if a proposition in S follows from a non-overruled argument in *Args*(*A*) which has p as its premise.

Thus, if a proposition in the explanation follows from a source of evidence the explanation covers that source of evidence. So if, for example, an explanation S covers five pieces of evidence, then $ec_S = 5$. Note that here we do not intend to define an objective probabilistic measure for the quality of stories; instead the notion of evidential coverage aids us in comparing explanations, viz.: an explanation S is better than an explanation $S´$ if its evidential coverage is higher. Note how the above definition ensures the plausibility of the evidential generalizations: if an argument that links a certain piece of evidence to an explanation is overruled, that piece of evidence does not count towards the evidential coverage of an explanation.

The explananda, while they also follow from evidence and are also events in an explanation, are treated differently from other events in an explanation; explananda cannot be attacked and providing arguments from evidence for an explanandum does not increase an explanation's evidential coverage. This is because we do not want to reason about *what* should be explained but instead we want to reason about *how* certain events are explained. In section 3.2 this point will be made clearer.

Let us illustrate the combined framework with a simple example, adapted from Wagenaar et al. ([19], page 35). The example concerns the Haaknat case, in which a supermarket was robbed. The police conducted a search operation in a park near the supermarket, hoping to find the robber. Haaknat was found hiding in a moat in the park and the police, believing that Haaknat was the robber, apprehended him. Haaknat, however, argued that he was hiding in the moat because earlier that day, he had an

argument with a man called Benny over some money. According to Haaknat, Benny drew a knife so Haaknat fled and hid himself in the moat where the police found him. The explanandum in this case is 'Haaknat is found by the police'. In figure 1 the two explanations for this explanandum are represented in a simple graph.
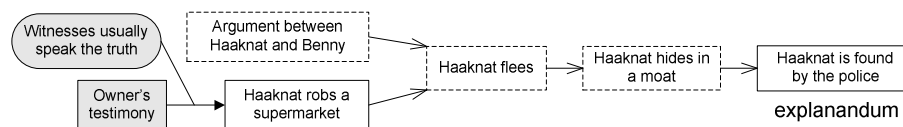

Figure 1: two explanations for the fact that Haaknat was found in the moat

In the figure, the causal theories combined with the hypotheses are represented as white boxes, where the variables in the causal theories have been instantiated with the constants from the hypotheses. Causal relations are rendered as arrows with an open head. A piece of evidence, namely, that the owner of the supermarket testified that it was Haaknat who robbed his shop, is represented as a grey box and the evidential generalization is represented as a grey rounded box; evidential relations are rendered as arrows with a closed head. In the example 'Haaknat robs a supermarket' follows from a non-overruled evidential argument. This can be seen in the figure, where events that are not supported by evidence are in a dotted box and events that are supported by evidence in a box with a solid line. The explanandum of course also follows from evidence (in this case a police report and Haaknat's own testimony). However, the links between this evidence and the explanandum have not been rendered in the above figure, because we want to focus on whether the explanation is supported by evidence and not on whether the explanandum is supported by evidence. The bottom explanation (that Haaknat robbed the supermarket) can be regarded as the best explanation because it has an evidential coverage of 1 while the other explanation (that Haaknat had an argument with Benny) has an evidential coverage of 0.

Both the argumentative and the IBE approach have been used separately to model evidential reasoning (see [4] for an example of a purely argumentative approach and [19] for an example of an IBE approach). We now briefly explain why we combine the two approaches instead of adopting just one of them.

A disadvantage of an argumentative approach is that it does not provide a complete overview of the case, as the original stories about 'what happened' are cut into pieces to become conclusions of different arguments and counter-arguments. An approach where stories are represented as causal networks and thus the overview of the case is retained is closer to how legal decision makers and investigators actually think about a case ([13],[15]). This was informally confirmed in our contacts with police detectives and lecturers of the Dutch police academy, in which we learned that crime investigators often visualise time lines and causal structures to make sense of a body of evidence.

However, a problem of a purely IBE approach is that sources of evidence such as testimonies are modelled as effects caused by the event for which they serve as evidence. The advantage of adding evidential arguments to the IBE approach is that in the combined theory, reasoning with sources of evidence is arguably more natural: events are inferred from evidence using evidential generalizations. In our informal contacts with the Dutch police we found that this is how crime analysts usually connect the available evidence with their temporal and causal models of a case.

Another problem of the IBE approach as it is usually modelled is that it is impossible to reason *about* the causal generalizations used in an explanation. In legal settings this is a limitation since it is well-known that in criminal cases the quality of

the generalisations used by crime investigators or lawyers to build explanations cannot be taken for granted ([1], [19]).

3. The dialogue game

The analysis of stories and evidence in a case is a process; exactly how this process takes form depends on who performs the analysis and the specific legal context the analysis is performed in. In a decision-making context, for example, the defence is confronted with a complete story about what happened, namely the prosecution's story. Usually, this story is already supported by evidence and the defence will try to attack the prosecutor's evidential arguments (by arguing, for example, that a witness is not trustworthy) or the defence gives an alternative explanation (for example, that an alleged killer acted in self-defence). In an investigation context, however, things are different. Often, a team of criminal investigators is faced with some initial evidence and they construct several possible stories (or scenarios) and then try to find new evidence that supports or discredits these scenarios. During the investigation there is constant interaction between the scenarios and the evidence: a scenario provides a frame in which new evidence can be interpreted and, at the same time, new evidence is used to support or discredit a scenario or to extend a scenario [15].

In this paper, we aim to model the dynamics of the process of analysing stories and arguments in a formal dialogue game, with which it should be possible to build, critically analyse and change explanations and their supporting arguments.

Dialogue games formulate principles for coherent dialogue between two or more players, and this coherence depends on the goal of a dialogue. In our previous work on dialogue games [3], one of the players made a claim which he had to defend, while the other player's goal was to dispute this claim. The goal of the dialogue game was to resolve this difference of opinion in a fair and effective way. By contrast, the current dialogue game is meant to regulate a discussion between analysts in a criminal case. In such a setting the players have identical roles since they both want to find a plausible and evidentially well-supported explanation for the explananda. Moreover, none of the players really wants to win, since they have the joint goal to find the best explanation of the explananda. As explained in the introduction, our dialogue game is designed to promote this joint goal of the players by forcing them in an adversarial setting, where they technically have the aim to 'win', so that all sides of a case are explored. Accordingly, the game allows both players, given an initial body of evidence, to propose, criticise and defend alternative explanations for what happened. The idea behind enforcing such an adversarial setting is to avoid the well-known problem of 'tunnel-vision' or confirmation bias, where one explanation is taken as the right one and the investigation focuses on finding evidence that supports this explanation while dismissing evidence that contradicts this explanation. Note that while the game has two players, extending the dialogue game to accommodate for more players is easy, thus allowing our dialogue game to support discussions between groups of analysts.

Now, in a dialogue the players build a framework for evidential reasoning *ER* by performing speech acts from a communication language L_c. With these speech acts, explanations can be given for the explananda *F*, and arguments can be moved for supporting explanations or for attacking explanations or other arguments, thus continually updating the framework *ER*. One part of the dialogue game is a *protocol*, which specifies the allowed moves at a certain point in the dialogue. Such a protocol is

essentially a normative model for how the process of an analysis of evidence and explanations should take place.

The dialogue game also has *commitment rules*, which specify the effects of a speech act on the propositional commitments of the dialogue participants. For instance, explaining the explananda with an explanation commits the speaker to the explanation and retracting a previously moved argument removes this argument from the speaker's commitments. Commitments can be used to constrain the allowed moves, for example, to disallow moves that make the speaker's commitments inconsistent. They can also be used to define *termination* and *outcome* of a dialogue. Recall that the objective of the game is to find the best explanation for the explananda so the outcome of a dialogue is an explanation together with its supporting arguments and the dialogue terminates if both players are committed to the best explanation. In addition, for nonterminated dialogues a notion of the *current winner* can be defined; this is the adversarial element of the dialogue. The current winner is the player that is committed to the currently best explanation for the explananda. The notion is used to control turn taking, with a rule that a player is to move until he has succeeded in becoming the current winner (cf. [9]).

We now turn to the definitions of the elements of our dialogue game. Because of space limitations, the definitions will in some places be semiformal. Below, $AR \in Args(A)$ and $\varphi \in wff(L)$, where L is the underlying logic of the framework (see section 2). Dialogues take place between two players, p_1 and p_2. The variable a ranges over the players, so that if a is one player, then $\bar{a}$ is the other player.

The communication language L_c consists of the following locutions or speech acts:

- *argue AR*. The speaker states an argument.
- *explain (E, S)*. The speaker provides an abductive explanation $S = H \cup T$ for a set of propositions E.
- *concede φ*. The speaker admits that proposition φ is the case.
- *retract φ*. The speaker declares that he is not committed (any more) to φ.

The speech act *explain* is new while the other locutions are well-known from the literature. A *dialogue d* is now a sequence of utterances of locutions from L_c, where d_0 denotes the empty dialogue. Each utterance is called a *move*. The speaker of a move m is denoted by $s(m)$.

3.1. Commitments

The players' commitments are influenced by the moves they do during a dialogue. At the start of a dialogue the commitments of both players consist of just the explananda from F. The set $Comms_s$ denoting the commitments of the speaker s is updated during a dialogue as follows. When s moves an *argue AR* move, the premises of AR and conclusions of AR are added to $Comms_s$; when s moves an *explain (E, S)* move, the elements from E and S are added to $Comms_s$; when s moves a *concede φ* move, φ is added to $Comms_s$ and when s moves a *retract φ* move, φ is deleted from $Comms_s$.

3.2. The framework in the dialogue protocol

Recall that in our setup the dialogue participants jointly build a framework for evidential reasoning *ER*. In this framework, the set explananda F is given and assumed nonempty and the players can update the framework by providing explanations or arguments using the speech acts. The set F does not change during the dialogue, so it

must be agreed upon before the dialogue starts. It is in theory possible to have an argumentative dialogue about what the explananda are. However, the purpose of the current dialogue is to find explanations for certain observations and to compare these explanations; a dialogue about what should be explained is a different kind of dialogue the details of which we leave for future research.

$ER(d) = ((H(d), T(d), F(d)), (G(d), I(d)))$ stands for the evidential reasoning framework after dialogue d. The elements of this framework also denote the elements after a certain dialogue d; so $H(d)$ is H after dialogue d, $T(d)$ is T after dialogue d etcetera. When the speaker s makes a move, the framework is updated as follows. When s moves an *argue AR* move, the generalizations in the argument AR are added to the set of evidential generalizations G and the other premises of AR are added to I. When s moves an *explain* $(E, (H' \cup T'))$ move, the hypotheses in H' are added to H and the causal generalizations in T' are added to T. When s moves a *retract* φ move and $\varphi \notin Comms_s(d)$, then φ is removed from its corresponding element in the framework.

3.3. Turn taking and winning

Before the protocol itself is defined, two related notions need to be defined, namely turn taking and winning. A player a is the *current winner* of a dialogue d if there is an explanation S, $S \in Expl(ER(d))$ and $S \subseteq Comms(a)$, and for each other explanation S', $S' \in Expl(ER(d))$ and $S' \neq S$, it holds that $ec_S > ec_{S'}$. So if there is only one explanation for the explananda in the current framework, a player is the current winner if he is the only player committed to that explanation. If there are more explanations, a player is the winner if he is the only player committed to the explanation that has the highest evidential coverage. Note that this definition of the current winner also allows that there is no current winner, namely when both players are committed to explanations with equal evidential coverage

With the notion of a current winner, a turn taking rule can be defined as follows: *Turn* is a function that for each dialogue returns the players-to-move, such that $Turn(d_0) = p_1$, $Turn(d, m) = \bar{a}$ if a currently wins d, else if there is no current winner and $Turn(d) = a$ then $Turn(d, m) = a$. Thus it is always the losing player's turn and even if he makes such a move that the other player is no longer the winner, he still has to become the winner himself. This situation ensures that both players try to advance and defend their respective explanations as opposed to the situation where one player gives an explanation and the other player constantly attacks this one explanation.

3.4. The protocol

The protocol P specifies the allowed moves at each stage of a dialogue. Its formal definition is as follows. For all moves m and dialogues d it holds that $m \in P(d)$ if and only if all of the following conditions are satisfied:

1. $Turn(d) = s(m)$
2. m was not already moved in d by the same player
3. $Comms_s(d, m) \nvdash \bot$
4. If m is an *argue AR* move (where φ is AR's conclusion), then $\varphi \notin F$ and
   - either $\varphi = \neg ge_i$ for some $ge_i \in G(d)$ or φ is a negation of an element in (I / I_E)
   - either $\varphi = \neg gc_i$ for some $gc_i \in T(d)$

- or $\exists S, S \in Expl(ER(d)), S \subseteq Comms(s)$ and $ec_S(d, m) > ec_S(d)$

5. If m is an *explain* (E, S) move, then
    - S explains E and $E \cap F \neq \varnothing$
    - S explains E and; $\forall e \in E: H(d) \cup T(d) \vdash e$ and e is a literal

6. If m is a *concede* φ move, then φ is in an element of $ER(d)$ and $\varphi \notin Comms_s(d)$

7. If m is a *retract* φ move, then $\varphi \notin F$ and $\varphi \in Comms_s(d)$

8. $\neg \exists S: S \in Expl(ER(d)), S \subseteq Comms_s(d), S \subseteq Comms_{\bar{s}}(d)$ and for each other explanation $S', S' \in Expl(ER(d))$ and $S' \neq S$, it holds that $ec_S > ec_{S'}$.

The first two conditions say that only the player-to-move can make allowed moves and that a player may not repeat his moves. Condition (3) regulates the players' logical consistency. The first point of condition (4) states that an argument may be moved if it attacks another argument, that is, it attacks an evidential generalization or it attacks an element from input that is not a source of evidence. The second point states that an argument may be moved if attacks an explanation, that is, it attacks a causal generalization. The third point states that an argument may be moved if it improves the evidential coverage of an explanation to which the player is committed. Condition (5) states that an *explain* move may be done if it explains (see definition on page 3) an explanandum, or if it explains a literal that follows from the current hypotheses and the current theory. Condition (6) ensures that a player concedes a proposition only if it is in the current framework and the player is not already committed to it. Condition (7) says that a player can only retract a proposition to which he is committed. Finally, condition (8) implies that a dialogue terminates if both players are fully committed to the best explanation.

In the current dialogue game, the legality of moves is defined in terms of the current framework for evidential reasoning: every move must be a sensible operation on *ER*. This way of defining the relevance of moves is different from, for example, [12], where relevance is enforced by a strict protocol and [16], where relevance is enforced by the reply structure on the speech acts in the communication language.

4. Example

For the example we return to the Haaknat case on page 5. The set of explananda F in this case is {Haaknat is found by police}. Player p_1 starts the dialogue by providing an explanation for this explanandum:

p_1: *explain* ({Haaknat is found by police}, {Haaknat robs supermarket} $\cup T_1$)

      where $T_1 = \{$**gc$_1$**: x robs supermarket $\Rightarrow_C x$ flees, **gc$_2$**: x flees $\Rightarrow_C x$ hides in a moat,

                gc$_3$: x hides in moat $\Rightarrow_C x$ is found by police$\}$

Now p_1 is winning, because he is committed to the one explanation for F, which is obviously the best explanation. p_2 at this point only has one option if he wants to become the current winner: he has to provide an explanation for F which is better than p_1's explanation.

p_2: *explain* ({Haaknat is found by police}, {argument between Haaknat and Benny} $\cup T_2$

      where $T_2 = \{$**gc$_4$**: argument between x and $y \Rightarrow_C x$ flees, **gc$_2$**, **gc$_3$**$\}$

After providing this explanation, it is still p_2's turn, as the explanation he has provided is not better than p_1's explanation. p_2 supports his explanation by providing an argument. Below, $\gg$ stands for the application of the defeasible modus ponens.

p_2: *argue* AR_1:
> ($\mathbf{e_1}$: Haaknat's testimony "I had an argument with Benny" $\wedge$ $\mathbf{ge_1}$: witness testifies that "p" $\Rightarrow_E$ p)
> $\gg$ argument between Haaknat and Benny

Now p_2 is the current winner: there is one piece of evidence in the case and it is covered by p_2's explanation. The current framework is pictured in the figure below (the different arrows and boxes are explained on page 5). For each event, it is indicated which players are committed to that event.
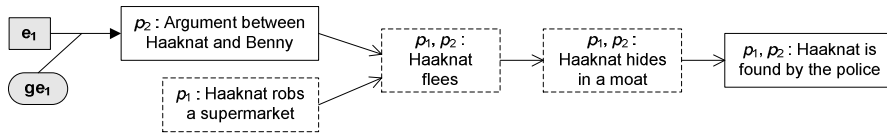
At his point p_1 can, for example, provide an argument for $\neg\mathbf{gc_4}$; if he does this, then p_2 no longer has an explanation for F so p_1 automatically has the best explanation. He can also try to support "Haaknat robs a supermarket" with at least two pieces of evidence, as this would make p_1's explanation have a higher evidential coverage. Another option is to decrease the evidential coverage of p_2's explanation by defeating the argument AR_1. Suppose that p_1 chooses to take this last option:

p_1: *argue* AR_2:

($\mathbf{e_2}$: Haaknat is a suspect in the case $\wedge$ $\mathbf{ge_2}$: suspects do not make reliable witnesses) $\gg$ $\neg\mathbf{ge_1}$

For the sake of the example, assume that this argument defeats AR_1. p_1 is still not the current winner: both explanations have an evidential coverage of 0, so p_1 has to make another move in order to make his explanation better or p_2's explanation worse. p_1 could increase the evidential contradiction of p_2's explanation by providing an argument for $\neg$(argument between Haaknat and Benny). However, in this case p_1 chooses to increase the evidential coverage of his own explanation. He does this by first expanding the explanation and then supporting it with evidence:

p_1: *explain* ({Haaknat robs supermarket}, {Haaknat is from Suriname} $\cup T_3$
> where $T_3 = \{\mathbf{gc_5}$: x is from Suriname $\Rightarrow_C x$ robs supermarkets$\}$

p_1: *argue* AR_3:
($\mathbf{e_1}$: Haaknat's birthplace is "Republic of Suriname" $\wedge$

  $\mathbf{ge_2}$: x birthplace is "Republic of Suriname" $\Rightarrow_E x$ is from Suriname) $\gg$ Haaknat is from Suriname

The following picture represents the current situation. The light grey argumentation arrow means that the inference is defeated and the arrow connected to AR_2 stands for a defeat relation.
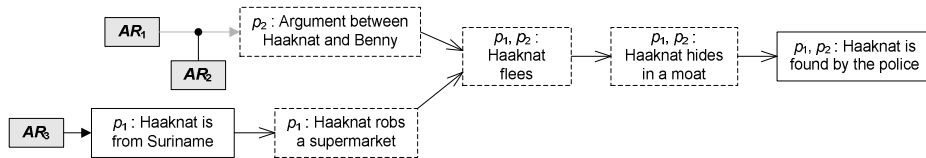

p_1 is now the winner: his explanation has an evidential coverage of 1 while p_2's evidential coverage is 0. However, part of p_1's generalization is based on the

generalization "people from Suriname rob supermarkets". p_2 does not agree with this and he argues that the generalization is based on prejudice:

p_2: *argue* AR_4: ($\mathbf{i_1}$: $\mathbf{gc_5}$ is based on prejudice, $\mathbf{ge_2}$: $\mathbf{gc_i}$ is based on prejudice $\Rightarrow_E \neg\mathbf{gc_i}$) $\gg \neg\mathbf{gc_5}$

Note that AR_4 is not based on evidence, so it is possible for p_1 to attack $\mathbf{i_1}$. By attacking $\mathbf{gc_5}$, p_2 ensures that {Haaknat is from Suriname} $\cup$ T_1 $\cup$ T_3 is no longer an explanation. This is shown in the following figure, where AR_4 attacks the causal generalization and the part of p_2's explanation that is no longer considered is rendered in light grey.
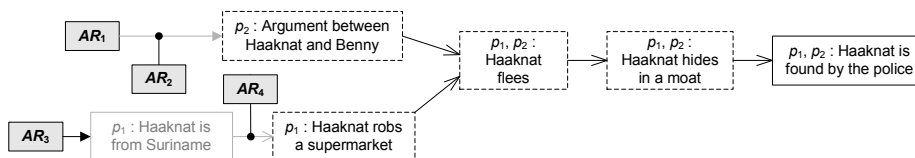
p_2's explanation {argument between Haaknat and Benny} $\cup$ T_2 and p_1's explanation {Haaknat robs supermarket} $\cup$ T_1 now both have an evidential coverage of 0, so p_2 needs to make another move in order to become the winner.

5. Conclusions

In this paper we have shown how the combined story and argumentative approach to reasoning with evidence can be fit into a formal dialogue game. This dialogue game not only allows for the construction of defeasible arguments but also allows the players to explain events using abductive explanations, which can then be compared according to their conformity with the evidence in the case. Furthermore, the argumentative part of the game and framework allows players to have critical discussions about the plausibility and validity of the causal and evidential generalizations. The winning and turn taking conditions ensure that the players are forced to advance and improve their own explanations as well as criticize their opponent's explanations, which hopefully avoids 'tunnel-vision' or confirmation bias.

Our dialogue game can be seen as a guideline for a critical discussion between investigators in a criminal case. Furthermore, we intend to use the ideas presented in this paper in the further development of our software for police investigators, *AVERs*, which is currently being developed by other members of our project [7].

The precise form of our dialogue game is, to our knowledge, new. As far as we know, there are only two other dialogue games that allow the players to jointly build a theory ([11],[3]) build a Bayesian network and an argumentation-graph, respectively, whereas in our game a combined framework for argumentation and IBE is built. Our dialogue is a combination of an enquiry and a persuasion dialogue [21]; on the one hand, the players have a shared 'quest for knowledge' but on the other hand, the players try to persuade each other that their explanation is the best.

Acknowledgements

References

[1] L. Amgoud, L. Bodenstaff, M. Caminada, P. McBurney, S. Parsons, H. Prakken, J. van Veenen, G. Vreeswijk, Final review and report on formal argumentation system. ASPIC Deliverable D2.6, 2006

[2] T.J. Anderson, D.A. Schum and W.L. Twining. *Analysis of Evidence, 2nd edition*. Cambridge University Press, 2005.

[3] F.J. Bex and H. Prakken. Reinterpreting arguments in dialogue: an application to evidential reasoning. *Proceedings of JURIX 2004*, pp.119 – 129. IOS Press, 2004.

[4] F.J. Bex, H. Prakken, C. Reed and D. Walton, Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, pp. 125 – 165, 2003.

[5] F.J. Bex, H. Prakken and B. Verheij. Formalizing argumentative story-based analysis of evidence. *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, ACM Press, 2007.

[6] F.J. Bex, S.W. van den Braak, H. van Oostendorp, H. Prakken, H.B. Verheij & G.A.W. Vreeswijk Sense-making software for crime investigation: how to combine stories and arguments? *To appear in Law, Probability & Risk*, Oxford University Press, 2008

[7] S.W. van den Braak, G.A.W. Vreeswijk, and H. Prakken, AVERs: An argument visualization tool for representing stories about evidence. *Proceedings of the 11th International Conference on Artificial Intelligence and Law*. ACM Press, 2007.

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

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

[10] P. Lucas, Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review* 12, pp. 109 – 146, 1997

[11] S.H. Nielsen and S. Parsons, An Application of Formal Argumentation: Fusing Bayes Nets in MAS. *Computational Models of Argument: Proceedings of COMMA 2006*, IOS Press, 2006

[12] S. Parsons, M. Wooldridge and L. Amgoud, An analysis of formal inter-agent dialogues *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 394-401, ACM Press, 2002.

[13] N. Pennington and R. Hastie. The story model for juror decision making. In R. Hastie (eds.), *Inside the Juror, The Psychology of Juror Decision Making*. Cambridge University Press, 1993.

[14] J. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, 1995.

[15] C.J. de Poot, R.J. Bokhorst, P.J. van Koppen and E.R. Muller, *Rechercheportret - Over dilemma's in de opsporing* (*Investigation portrait – about dilemmas in investigation*), Kluwer, 2004.

[16] H. Prakken, Coherence and flexibility in dialogue games for argumentation. *Journal of Logic and Computation* 15, pp. 1009-1040, 2005.

[17] H. Prakken and G. Sartor, Argument-based extended logic programming with defeasible priorities. *Journal of Applied Non-classical Logics* 7, pp.25 – 75, 1997.

[18] Schank, R. C. and Abelson, R. P. (1977) *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*, Lawrence Erlbaum, Hillsdale, NJ.

[19] P. Thagard, Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks. *Applied Artificial Intelligence*, 18, 231-249, 2004.

[20] W.A. Wagenaar, P.J. van Koppen and H.F.M. Crombag, *Anchored Narratives. The Psychology of Criminal Evidence*. Harvester Wheatsheaf, 1993.

[21] D.N. Walton and E.C.W. Krabbe, *Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning*. State University of New York Press, 1995.