Towards an integrated theory of causal scenarios and evidential arguments

Floris BEX
Department of Information and Computing Sciences, Utrecht University

Abstract. The process of proof is one of inference to the best explanation, in which alternative scenarios are supported and attacked by arguments. This combination of scenarios and arguments was previously presented as a formal hybrid theory. In this paper, the aim is to further integrate scenarios and arguments by defining a notion of attack between alternative explanations. Thus, scenarios and arguments can be incorporated in the same dialectical framework.

Keywords. causal reasoning, evidential reasoning, scenarios, arguments

1. Introduction

The study of reasoning with evidence constitutes the study of the process of proof, which involves reasoning with evidence and commonsense knowledge in order to establish whether something is or was the case. Whether in law, medicine or science, the reasoning in the process of proof can be viewed as *inference to the best explanation* (IBE): alternative scenarios that explain the evidence are constructed and ultimately the best scenario is chosen.

The logical frameworks for IBE that have been proposed in the AI literature [7][10] use causal rules to model the scenarios, which are then compared on basic criteria such as the minimum number of hypothetical assumptions. Because these frameworks were originally intended for automated diagnosis within relatively small and pre-defined domains, they do not incorporate aspects like reasoning with both causal rules (*c causes e*) and evidential rules (*e causes c*), reasoning about the causal rules that comprise the scenarios and more fine-grained scenario comparison criteria. Hence, these frameworks are less suited to modelling and supporting more complex tasks such as analysing and making sense of the evidence in large criminal cases [4][13].

One way to expand the analytical possibilities of these classic frameworks for IBE is to combine them with a general argumentation framework for evidential reasoning, so that arguments based on evidence can be used to support and attack the scenarios. This is the approach taken by the *hybrid theory* [4], in which causal, scenario-driven reasoning is combined with argumentative, evidence-driven reasoning.

The expressive power of the hybrid theory, however, negatively influences its formal rigour. Because of the combination of two separate frameworks, causal scenarios are compared differently from evidential arguments. For example, two contradictory arguments attack each other but two contradictory scenarios do not, and two alternative scenarios are in competition, whilst two arguments that present alternative explanations are not. Furthermore, the status of arguments is determined using standard semantics of [8], whereas the status of scenarios is determined with more absolute measures (e.g., the more observations a scenario explains the better it is [11][13]). Whether or not a certain claim is considered acceptable can therefore depend on if it is modelled as part of a scenario or an argument. What is missing from the hybrid theory is a standard way to compare both alternative explanations and contradictory claims irrespective of whether they are part of a scenario or an argument.

The present paper aims to further integrate arguments and scenarios by treating them as equals in the dialectical process. It will be shown how alternative scenarios can be considered as counterarguments. Furthermore, it will be discussed how arguments that present alternatives can attack each other. Thus, the status of both scenarios and arguments can be determined using the semantics of [8]. Note that these amendments do not lead to a loss of expressiveness: it is still possible to reason about the scenarios and the original criteria for comparing scenarios are maintained.

2. A Hybrid Theory of Scenarios and Arguments

Two influential methods for the analysis of evidence are the argument-based approach [2], which focuses on *arguments* based on evidence, and the story-based approach [11], which uses hypothetical stories or *scenarios* to explain the evidence. In the argument-based approach the reasoning goes from evidence to conclusions by the application of evidential rules (*e is evidence for c*), and arguments for and against the conclusion are considered and compared. In the scenario-based approach multiple scenarios consisting of causal rules (*c is a cause for e*) are constructed to explain the evidence, and the scenarios are considered and compared.

Both the evidential argument and the causal scenario approach have advantages. An inference from a piece of evidence (e.g. a witness testimony) to a conclusion is best captured using evidential rules [4][6], and arguments based on evidence can be compared in a solid but intuitive framework [8]. However, certain parts of a case (such as the cause of death, or the behaviour of the suspect) are more easily captured using causal rules [6]. Furthermore, scenarios help in providing an overview of the case as they are understandable stories that explain the evidence in a natural way [11].

The combination of evidential arguments and causal scenarios seems to be an intuitive and analytically useful perspective for looking at cases and evidence. Hence, in [4] a hybrid theory for reasoning with arguments and scenarios is proposed, a combination of a formal theory of causal-abductive reasoning *CT* based on classical model-based approaches [7][10] and an evidential argumentation theory *ET* based on the ASPIC+ framework [12]. The logical language of this theory contains a connective $\Rightarrow$ (defeasible implication), and the logic includes an inference rule for this connective (defeasible modus ponens). Rules can be either causal ($p \Rightarrow_C q$) or evidential ($q \Rightarrow_E p$).

A causal theory *CT* contains a set of evidence, a set of causal rules and a set of hypotheticals (literals). We can then construct *scenarios*, sequences of hypotheticals, rules and elements that are derived from some previous element by application of defeasible modus ponens. For example, consider S_1 [9].

S_1: [rain, rain $\Rightarrow_C$ grass_wet, grass_wet]

Here, rain is hypothesized to explain the evidence grass_wet. A scenario *S explains* an observation $e \in E$ iff $S \vdash_C e$, where $\vdash_C$ stands for logical consequence according to the set of common deductive inference rules extended with modus ponens for $\Rightarrow_C$. For example, scenario S_1 explains grass_wet. An important part of abductive reasoning is the consideration of alternative scenarios that explain the evidence. For example, an alternative explanation for grass_wet is that the sprinkler was on during the night.

S_2: [sprinkler_on_, sprinkler_on $\Rightarrow_C$ grass_wet, grass_wet]

We can also infer conclusions from the evidence using evidential rules, thus creating arguments that can be used to support or attack scenarios.

In the evidential theory *ET* we have a set of evidence and a set of evidential rules. By applying defeasible modus ponens, we can then build arguments based on evidence that support or attack scenarios. Take, for example, argument A_1.

A_1: [grass_looks_wet, grass_looks_wet $\Rightarrow_E$ grass_wet, grass_wet]

Arguments can *rebut* each other (opposite conclusion) and one argument can *undercut* another argument (attacking the inference) [12]. Given a collection of arguments and their attack relations, the status of the arguments can be determined according to [8].

In the hybrid theory, the status of alternative scenarios depends on various criteria. One example is *evidential support*, the number of pieces of evidence that support a scenario (*i.e.*, evidence that is part of the scenario or is the premise of an argument that has as its conclusion an element in the scenario). The second is *evidential contradiction*, the number of pieces of evidence that contradict a scenario (*i.e.*, that are premises of an argument that contradicts the scenario). These criteria can be used to compare scenarios, and by contrasting them to the total number of pieces of evidence in a case we get an indication of the relevance and coverage of a scenario [11].

Causal and evidential reasoning are closely intertwined: if we have a causal rule *c causes e*, and *c* is a normal ('default') cause for *e*, then we will usually also accept that *e is evidence for c*. For example, rain can cause the grass to be wet so the observation of wet grass can be seen as evidence for the fact that it rained. Exactly how we should model a relation (causally or evidentially) seems arbitrary. [6] found that while in some situations, people quite consistently chose either causal or evidential, there are also many examples where people interpret the relation evidentially or causally in equal measure. Furthermore, some aspects of a case will more conveniently be modelled in terms of evidential arguments, others in terms of causal scenarios.

Under the right assumptions, evidential and causal reasoning are formally equivalent [10]: either we use causal rules and perform abductive reasoning or we use evidential rules and perform modus ponens style reasoning. So, in a formal theory, we should be able to change an evidential link into a causal link and still get the same result from the scenario and argument comparison. For the hybrid theory, this is not the case. Take, for example, the two alternative scenarios S_1 and S_2. If we change the causal rules into the reverse evidential rules, we get two arguments, viz.

A_2: [grass_wet, grass_wet $\Rightarrow_E$ rain, rain]
A_3: [grass_wet, grass_wet $\Rightarrow_E$ sprinkler_on, sprinkler_on]

These two arguments, however, do not attack each other, while the two scenarios S_1 and S_2 were in competition. Similarly, say that we have the following argument.

A_4: [$\neg$trees_wet, $\neg$trees_wet $\Rightarrow_E$ $\neg$rain, $\neg$rain]

This argument will contradict S_2, adding to its evidential contradiction. A_4 will also attack A_2, but since determining the 'winning' argument is done differently than determining the 'winning' scenario, it is not guaranteed that a case with S_1 and A_4 has the same outcome as a case with A_2 and A_4. Also, if we turn A_4 into a scenario:

S_3: [¬rain, ¬rain $\Rightarrow_C$ ¬trees_wet, ¬trees_wet]

This scenario does not attack S_1, even though the S_1 and S_2 are clearly incompatible. Thus we cannot capture the idea of an alibi scenario, which is not an alternative explanation (it does not explain the main questions in a case, e.g. why the victim died) but it does contradict the prosecution's scenario (that the suspect killed the victim).

3. An integrated formal framework

The above discussion indicates that we need to reconsider the relation between arguments and scenarios: we need to be able to reason with alternative explanations even if they are modelled as arguments, using evidential rules, and we need to be able to consider contradictory causal scenarios. The comparison of alternative scenarios and attacking arguments should be drawn level so that we have a fully integrated theory instead of a hybrid of causal, scenario-based and evidential argumentative reasoning.

The logical language and logic of the integrated theory *IT* are the same as for the hybrid theory: there are connectives $\Rightarrow_E$ and $\Rightarrow_C$ for evidential and causal rules, and a defeasible modus ponens (DMP) inference rule for these connectives. Like in the hybrid theory, there are sets of causal and evidential rules.

Definition [Rules] *The set of rules is* $R = R_E \cup R_C$, *where* R_E *is the set of evidential rules and* R_C *is the set of causal rules.*

Following ASPIC+ [12], there is one knowledge base (with appropriate subsets).

Definition [Knowledge base] *A knowledge base is a set* $K = K_e \cup K_h$, *where* K_e *is the set of evidence and* K_h *is the set of hypotheticals.*

In the hybrid theory, arguments and scenarios are similar in structure: both are derivations in a defeasible logic. Scenarios can be considered as (hypothetical) causal arguments, which puts them at the same level as evidential arguments.

Definition [Arguments] *A* causal/evidential argument *is a finite sequence* $[\varphi_1, \ldots, \varphi_n]$, *where* $n > 0$*, such that for all* φ_i $(1 \leq i \leq n)$:

- $\varphi_i \in K \cup R_C$ (*causal argument/scenario*), $\varphi_i \in K \cup R_E$ (*evidential argument*); *or*
- φ_i *follows from* $\psi_1, \ldots, \psi_n \in \{\varphi_1, \ldots, \varphi_{i-1}\}$ *by application of DMP.*

We say that $\mathrm{Conc}(A) = \{\varphi \in A \mid \varphi$ *is the last element in the sequence*$\}$ *is the conclusion of argument* A, *and argument* B *is a (proper) subargument of* A *iff* B *is a (proper) subsequence of* A *that is also an argument*.

In our example, $A_1 – A_4$ are evidential arguments and $S_1 – S_3$ are causal arguments. We can now define when two arguments are alternatives.

Definition [Alternative Causal Arguments] *Causal argument* C_1 *is an alternative to causal argument* C_2 *if* $\mathrm{Conc}(C_1) = \mathrm{Conc}(C_2)$ *and there exist* C_1', C_2' *such that* $\mathrm{Conc}(C_1') \neq \mathrm{Conc}(C_2')$ *and* C_1', C_2' *are proper subarguments of* C_1,C_2, *respectively*.

The idea here is that there are two possible causes, Conc(C_1') and Conc(C_2'), for the effect denoted by the same conclusion of C_1 and C_2. In our example, S_1 and S_2 are alternatives: they both have the conclusion shoes_wet, but S_1 has a subargument with conclusion rain and S_2 has a subargument with conclusion sprinkler_on.

Evidential arguments can also be alternatives: an evidential rule is essentially an explicit expression of an abductive reasoning step applied to a causal rule.

Definition [Alternative Evidential Arguments] *Evidential argument* E_1 *is an alternative to evidential argument* E_2 *if* Conc(E_1) ≠ Conc(E_2) *and there exist* E_1', E_2' *such that* Conc(E_1') = Conc(E_2') *and* E_1', E_2' *are proper subarguments of* E_1, E_2, *respectively*.

This definition, where the conclusion of an argument is different but some of its premises are the same, is similar to [1], who discuss how arguments that represent different actions that lead to the same goal can be seen as competing alternatives. The idea is again that there are two possible causes, Conc(E_1) and Conc(E_2), for one effect (Conc(E_1') and Conc(E_2')). A_2 and A_3 are alternatives: A_2 implies that grass_wet was caused by rain, whilst A_3 implies that grass_wet was caused by sprinkler_on.

Definition [Alternative Evidential/Causal Arguments] *Evidential argument* E *is an alternative to causal argument* C *if there exist* C', E', C'', E'' *such that* Conc(C') = Conc(E') *and* Conc(C'') ≠ Conc(E''), *where* C', E' *are subarguments of* C, E, *respectively, and* C'' *is a subargument of* C' *and* E' *is a subargument of* E''.

Again, the idea is that there are two alternative causes Conc(C'') and Conc(E'') for one and the same effect, Conc(C') = Conc(E'). As a clarification of the above definition, consider the different subarguments of the alternative arguments S_1 and A_3.

S_1': [rain, rain $\Rightarrow_C$ grass_wet, grass_wet]	S_1'': [rain]
A_3 / A_3'': [grass_wet, grass_wet $\Rightarrow_E$ sprinkler_on, sprinkler_on]	A_3': [grass_wet]

Conc(S_1') = Conc(A_3') and Conc(S_1'') ≠ Conc(A_3''), where S_1' and A_3' are subarguments of S_1 and A_3, S_1'' is a subargument of S_1' and A_3' is a subargument of A_3''.

Figure 1 visualizes the different alternative arguments. Arrows with open arrowheads are causal rules and arrows with closed arrowheads are evidential rules.

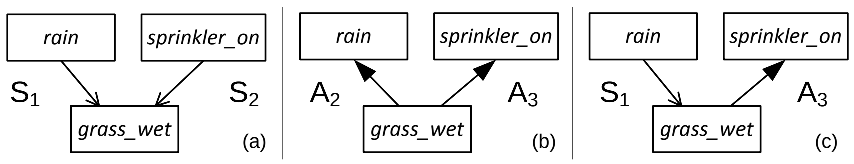

Figure 1: alternative causal (a), evidential (b) and causal/evidential (c) arguments

In explanatory reasoning from effect to cause the dialectical component lies in the alternatives: two explanations (whether scenarios or arguments) are in competition if they are alternative causes for the same consequence. Thus, we can use these definitions of alternatives in the definition of attack between arguments.

Definition [Attack] *Argument* A *attacks argument* B *iff*

- *There are two subarguments* A' *and* B' *that have an opposite conclusion, and the conclusion of* B' *is not in* K_e (*rebut*); *or*
- *There is a subargument* A' *which has conclusion* ¬(p $\Rightarrow_{E/C}$ q) *and there is an application of defeasible modus ponens to* p $\Rightarrow_{E/C}$ q *in* B (*undercut*); *or*
- A *and* B *are alternatives* (*alternative attack*).

Take the following examples of causal arguments (scenarios) and evidential arguments.

S_1: [rain, rain $\Rightarrow_C$ grass_wet, grass_wet]
S_2: [sprinkler_on_, sprinkler_on $\Rightarrow_C$ grass_wet, grass_wet]
S_3: [$\neg$rain, $\neg$rain $\Rightarrow_C$ $\neg$trees_wet, $\neg$trees_wet]
A_1: [grass_looks_wet, grass_looks_wet $\Rightarrow_E$ grass_wet, grass_wet]
A_5: [$\neg$trees_look_wet, $\neg$trees_look_wet $\Rightarrow_E$ $\neg$trees_wet, $\neg$trees_wet]

S_1 and S_3 mutually attack (rebut) and S_1 and S_2 also mutually attack (alternative attack).

Given arguments and attacks, it can be informative to consider particular combinations arguments representing the positions in a case.

Definition [Position] *Given an integrated theory* IT, *a position* P *is a conflict-free set of arguments such that for each evidential argument* A *in* P, *there is a subargument* A' *which has as its conclusion an element in a causal argument* S $\in$ *P*.

So a position represents a consistent outlook on a case consisting of causal arguments and the evidential arguments supporting them. In Figure 2, the arguments from the example are shown as an argumentation framework [8], where arrows stand for attack relations, with the two positions P_1 and P_2 superimposed on the argumentation framework. It can be seen that S_2, S_3 and A_4 form a coherent whole that is in conflict with S_1.

We could now define evidential support for a position: this would simply be the number of pieces of evidence that are the premise of an argument in that position. Evidential contradiction can be defined similarly. This then allows us to indicate preferences between positions (e.g. the position with the higher evidential support is preferred) and thus give a rich way of defining preferences between sets of arguments based on evidence (cf. [1]). In the example, P_2 (supported by A_1 and A_5) is the preferred position since it has a higher evidential support than P_1 (supported by A_1).

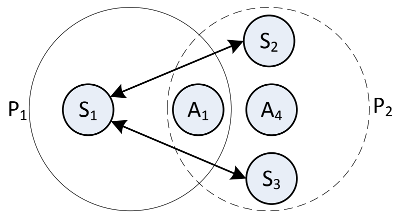

Figure 2: Two positions in a case with attacking arguments/scenarios

In the integrated theory we want to be able to reverse the direction of the rules, from causal to evidential and vice versa, without influencing the outcome of the case. However, causal and evidential rules – and more broadly, scenarios and arguments – behave as *communicating vessels* [5]: a change in an argument-oriented version of a case analysis requires a matching change in a story-oriented analysis. For example, if we change a causal rule $p \Rightarrow_C q$ into its corresponding evidential rule $q \Rightarrow_E p$ and there is no existing evidential argument with q in its conclusion, we have to add q to the set K, as otherwise we cannot derive p using the rule $q \Rightarrow_E p$. Furthermore, any arguments with conclusions $\neg(p \Rightarrow_C q)$ have to be changed into arguments with the conclusion $\neg(q \Rightarrow_E p)$, so that they will still undercut any evidential argument corresponding to the original causal argument that was undercut by $\neg(p \Rightarrow_C q)$.

If, however, we consistently apply these changes whenever we change a causal rule into an evidential rule or vice versa, the outcome of a case will stay the same despite the changes. In our example, say we change the causal rule rain $\Rightarrow_C$ grass_wet into the evidential rule grass_wet $\Rightarrow_E$ rain. This deletes the causal argument S_1 and gives us a new evidential argument A_1', viz.

A_1': [grass_looks_wet, grass_looks_wet $\Rightarrow_E$ grass_wet, grass_wet, grass_wet $\Rightarrow_E$ rain, rain]

Now rain and sprinkler_on will still be alternatives and S_2 and the new A_1' will attack each other. Similarly, S_3 will attack A_1' and thus still argue against rain. Furthermore, the positions in the case contain the same information: in P_1, S_1 is swapped for A_1' but this does not influence the conclusions we can draw from each of the positions or their evidential support.

4. Conclusion

The hybrid theory [4] provides a way to reason with scenarios, arguments and evidence. The integrated theory presented in this paper expands the hybrid theory and further integrates the reasoning with scenarios and arguments. By treating scenarios as causal-abductive arguments, they can be integrated into the dialectical framework of [12][8]. Furthermore, as long as we adhere to the principle of communicating vessels, we can freely exchange causal rules for their corresponding evidential rules and vice versa without affecting which conclusions we can draw in the case and the makeup of the positions in the case. Hence, the existing difficulties around whether or not to model a relation as causally or evidentially are alleviated. This approach is also advocated by [13], whose explanatory coherence theory does not include directional arrows but just links between ‘an explanation and what it explains’.

One distinction that has not been explicitly made in this paper is that between argumentation, where the truth of some claim is disputed and we want to show *that* the claim is true, and explanation, in which the claim is accepted but the question is *why* it is true. In fact, it could be argued that by equating scenarios with causal arguments, we fail to model explanations at all. However, it is the context of the reasoning that determines whether a particular reasoning structure (e.g. a causal sequence) is regarded as argument or explanation. The moment you present an alternative explanation, you are in fact arguing against the original explanation and thus the scenarios themselves can be treated as arguments. So the integrated theory does not dismiss explanation, but rather shows how argument and explanation are related.

Causal and evidential reasoning has been studied in depth by other authors. The most comprehensive work in AI is perhaps that in model-based diagnosis [7][10]. Because this work is originally intended to solve diagnosis problems with a pre-set causal theory (ellicitated from, for example, a medical expert), it largely avoids the problems associated with the sometimes vague distinctions between causal and evidential reasoning. Hence, these classic models are able to avoid complications reasoning with both causal and evidential rules and reasoning about rules. However, it has been argued [4][11][9][13] that one’s logical theory needs to be more expressive if we want to capture the subtleties of more broad commonsense reasoning as takes place in, for example, legal evidential reasoning. Pearl [11] makes a first step in combining causal and evidential reasoning in a principled way, but falls short of proposing a full-fledged logical theory. In future work we want to show that the integrated theory correctly adheres to the constraints Pearl proposes in his C-E framework.

An approach which has gained influence (at the expense of logical models such as [10][9]) is the Bayesian approach. In this approach the scenarios and evidence is a case are modelled as a Bayesian Network [16], a graphical representation of a joint probability distribution, where the nodes are variables (e.g., events, evidence) and the links represent the (in)dependencies between the variables. The directed links in a Bayesian network can be interpreted as causal rules, though a less strict criterion of 'correlation' is often also used. While Bayesian Networks are a powerful tool in evidential reasoning, a major shortcoming is that, in order to calculate the probability of the claims represented in the network, many conditional and prior probabilities are needed. More often than not these probabilities are simply not available. Furthermore, Bayesian Networks struggle with the same knowledge representation issues regarding causal and evidential reasoning as logical approaches do [15][16]. In my opinion, the work presented in this paper can therefore inform Bayesian approaches as much as logical approaches to evidential reasoning.

References

[1] Amgoud, L., and Vesic, S. (2013). Rich preference-based argumentation frameworks. *International Journal of Approximate Reasoning*, toappear.

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

[3] Bench-Capon, T.J.M. and Prakken, H. (2006) Justifying actions by accruing arguments. *Proceedings of COMMA-06*. Amsterdam, IOS Press, 247-258.

[4] Bex, F.J., Koppen, P.J. van, Prakken, H. and Verheij, B. (2010) A Hybrid Formal Theory of Arguments, Stories and Criminal Evidence. *Artificial Intelligence and Law* 18:2, 123-152.

[5] Bex, F.J. & Verheij, B. (2011) Solving a Murder Case by Asking Critical Questions: An Approach to Fact-Finding in Terms of Argumentation and Story Schemes. *Argumentation*, 26:3, 325-353.

[6] Braak, S.W. van den, Oostendorp, H. van, Prakken, H. and Vreeswijk, G. (2008) Representing narrative and testimonial knowledge in sense-making software for crime analysis. *Proceedings of JURIX 2008*. Amsterdam, IOS Press, 160-169.

[7] Console, L., and Torasso, P. (1991) A spectrum of logical definitions of model–based diagnosis. Computational Intelligence 7:3, 133 – 141.

[8] Dung, P.M. (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n–person games. *Artificial Intelligence* 77:2, 321 – 357.

[9] Pearl, J. (1988a) Embracing causality in default reasoning. *Artificial Intelligence* 35, 259 – 271.

[10] Poole, D. (1989). Explanation and Prediction: An Architecture for Default and Abductive Reasoning. Computational Intelligence, 5(2), 97–110.

[11] Pennington, N., and Hastie, R. (1993) Reasoning in explanation–based decision making. *Cognition* 49:1–2, 123 – 163.

[12] Prakken, H. (2010) An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1:93–124.

[13] Prakken, H. and Renooij, S. (2001) Reconstructing causal reasoning about evidence: a case study. Legal *Proceedings of JURIX 2001*, Amsterdam: IOS Press, 131-142.

[14] Timmer, S., Meyer, J.-J.Ch., Prakken, H., Renooij, S. and Verheij, B., Inference and Attack in Bayesian networks. *Proceedings of BNAIC 2013*, 199-206

[15] Thagard, P. (2004) Causal Inference in Legal Decision Making: Explanatory Coherence vs. Bayesian Networks. *Applied Artificial Intelligence* 18:3, 231 – 249.

[16] Vlek, C., Prakken, H., Renooij, S. and Verheij, B. (2013) Modeling Crime Scenarios in a Bayesian Network. *Proceedings of ICAIL 2013*, ACM Press, 150 -159.