Deductive and abductive argumentation based on information graphs

Remi Wieten [a,*], Floris Bex [a,b], Henry Prakken [a,c] and Silja Renooij [a]

[a] *Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC, Utrecht, The Netherlands*

[b] *Institute for Law, Technology and Society, Tilburg University, Warandelaan 2, 5037 AB, Tilburg, The Netherlands*

[c] *Faculty of Law, University of Groningen, Oude Kijk in 't Jatstraat 26, 9712 EK, Groningen, The Netherlands*

Abstract. In this paper, we propose an argumentation formalism that allows for both deductive and abductive argumentation, where 'deduction' is used as an umbrella term for both defeasible and strict 'forward' inference. Our formalism is based on an extended version of our previously proposed information graph (IG) formalism, which provides a precise account of the interplay between deductive and abductive inference and causal and evidential information. In the current version, we consider additional types of information such as abstractions which allow domain experts to be more expressive in stating their knowledge, where we identify and impose constraints on the types of inferences that may be performed with the different types of information. A new notion of attack is defined that captures a crucial aspect of abductive reasoning, namely that of competition between abductively inferred alternative explanations. Our argumentation formalism generates an abstract argumentation framework and thus allows arguments to be formally evaluated. We prove that instantiations of our argumentation formalism satisfy key rationality postulates.

Keywords: Argumentation, deduction, abduction, causal and evidential information, default reasoning

1. Introduction

In the legal and forensic domains, reasoning about evidence plays a central role in the rational process of proof [2, 4]. To aid in this process, various graph-based tools exist that allow domain experts to make sense of a mass of evidence in a case, such as mind maps [24, 36], argument diagrams [6, 24] and Wigmore charts [40]. Because of their informal nature, these tools typically do not directly allow for formal evaluation using AI techniques such as computational argumentation [14]. Hence, we wish to formalise and disambiguate analyses performed using such tools in a manner that (1) allows for formal evaluation and that (2) adheres to principles from the literature on reasoning about evidence [2, 4, 18, 26], while (3) allowing inference to be performed and visualised in a manner that is closely related to the way inference is performed and visualised by domain experts using such tools.

As we described in previous work [39], principles from the literature on reasoning about evidence state that inference is often performed using domain-specific *generalisations* [2, 4, 6], also called defaults [26, 33], which capture knowledge about the world in conditional form. A distinction can be made between *causal* generalisations (e.g. *'fire typically causes smoke'*) and *evidential* generalisations (e.g. *'smoke is evidence for fire'*) [4, 26]. In the current paper, we also consider generalisations that are neither causal nor evidential; examples are abstractions [4, 13] and mere statistical correlations. Inference can be performed

*Corresponding author. Email address: g.m.wieten@uu.nl. Telephone number: +31 6 3956 8028.

in a *deductive* or forward fashion, where from a generalisation (e.g. *'fire typically causes smoke'*) and its antecedent (fire), the consequent (smoke) is strictly or defeasibly inferred, and in an *abductive* [13, 18] or backward fashion, where from a causal generalisation or an abstraction and by affirming the consequent (smoke), the antecedent (fire) is defeasibly inferred. Note that the term 'deduction' is not consistently used in the literature, as it can either mean strict inference, in which the consequent universally holds given the antecedents (e.g. [21]) or defeasible inference, in which the consequent tentatively holds given the antecedents (e.g. [34]). To cover both meanings, in this paper 'deduction' is used as an umbrella term for both defeasible 'forward' inference and strict 'forward' inference.

Pearl [26, p. 264] argued that people generally consider it difficult to express knowledge using only causal generalisations, and in an empirical study, van den Braak and colleagues [9] found that while there are situations in which subjects consistently choose either causal or evidential modelling techniques, there are also many examples in which people use both types of generalisations in their reasoning. For instance, subjects often considered testimonies to be evidential, whereas a motive for committing an act is considered a cause for committing that act. This discussion illustrates that in formal accounts of reasoning about evidence, it is important to allow for causal and evidential generalisations [4]. Moreover, in this paper we show that it is important to also allow for abstractions and other types of generalisations, as these allow domain experts to be more expressive in stating their knowledge. The need for including these types of generalisations will become apparent from the examples we consider and the conceptional analysis of reasoning about evidence we provide.

When performing analyses using aforementioned tools such as mind maps, domain experts naturally mix the different types of generalisations and perform both deductive and abductive inferences, where the used generalisations and the inference type (deduction, abduction) are typically left implicit. Hence, in previous work [39] we set out to formalise analyses performed using these tools by providing a precise account of the interplay between the different types of inferences and generalisations and the constraints on performing inference we need to impose in terms of the *information graph* (IG) formalism. In this paper, we propose an extension of the IG-formalism, where in addition to causal and evidential generalisations we now also allow for abstractions and introduce a category of generalisations termed 'other', consisting of generalisations that are neither causal nor evidential nor abstraction such as aforementioned mere correlations, thereby increasing the expressivity of the IG-formalism. We particularly focus on identifying conditions under which performing inference with abstractions can lead to undesirable results. Specifically, care should be taken that no version of an event at a lower level of abstraction is inferred if an alternative version of this event at a lower level of abstraction was already previously inferred. Hence, we extend on the constraints imposed by Pearl's C-E system [26] which say that, in performing inference, care should be taken that no cause for an effect is inferred in case an alternative cause for this effect was already previously inferred. Moreover, in this paper we also consider exceptional circumstances under which the constraints of Pearl's C-E system should not be imposed, namely in case *enabling conditions* [12] are provided under which a generalisation may be used in performing inference. Based on these constraints and our conceptional analysis of reasoning about evidence, we define how deductive and abductive inference may be performed with IGs. Most existing formalisms that allow both inference types with causal and evidential information, abstractions, and other types of information are logic-based (e.g. [4, 5, 13, 21]); instead, we opt for a graph-based formalism to remain closely related to the way analyses are visualised using aforementioned graph-based tools.

The information specified in an IG serves as a source of information that can be used to facilitate the construction of AI systems for which formal semantics are defined. In earlier work [39], we investigated the application of our IG-formalism in facilitating the construction of Bayesian networks (BNs) [17], graphical models of joint probability distributions. In this paper, we instead focus on argumentation, where we propose an argumentation formalism based on IGs that allows for both deductive and abductive argumentation [38]. Previous work on abduction includes work on formal logical models of abductive reasoning (e.g. [13, 18]) and the work of Kakas and colleagues on abductive logic programming [20]. However, to the best of our knowledge, our proposed formalism is one of the first formalisms that models combined abductive and deductive reasoning in a formalism for structured argumentation. The closest to the current paper is Bex's integrated theory of causal and evidential arguments [5], which is based on the ASPIC$^+$ framework [21]. In Bex's integrated theory, the roles of generalisation and inference are not separated; instead, causal and evidential inferences are defined and arguments are constructed by forward chaining such inferences. In contrast to [5] we put special emphasis on the constraints that need to be imposed on the types of inferences that may be performed with the different types of generalisations, where we formally prove that arguments based on IGs indeed adhere to the identified constraints. Finally, compared to the ASPIC$^+$ framework [21], which only allows for deductive reasoning, we allow for both deductive and abductive reasoning and introduce a new type of conflict, namely conflict between competing alternative explanations [18], which is currently not accounted for in that framework. The relations to existing formalisms is further discussed in Section 6.

Our approach generates an abstract argumentation framework as in Dung [14], that is, a set of arguments with a binary attack relation, which thus allows arguments to be formally evaluated according to Dung's argumentation semantics. Besides allowing for rebuttal and undercutting attack, which are among the types of attacks that are typically distinguished in structured argumentation [21, 28], we also define the notion of alternative attack among arguments based on IGs, a concept based on the notion of competing alternative explanations that is inspired by [3, 5]. Alternative attack captures a crucial aspect of abductive reasoning, namely that of conflict between abductively inferred conclusions [18].

Our argumentation formalism extends a preliminary version proposed in [38] that was based on a more restricted version of our IG-formalism [39] in which only causal and evidential generalisations without enablers were considered. Moreover, in comparison to our earlier work [38] we now also prove that key rationality postulates [10] are satisfied by instantiations of our formalism, which implies that anomalous results as identified by [10] are avoided.

To summarise the main contributions of this paper, we propose an argumentation formalism that allows for both deductive and abductive argumentation, the latter of which has received relatively little attention in argumentation. Our argumentation formalism is based on an extended version of our IG-formalism, where in addition to causal and evidential generalisations we now also allow for abstractions and other types of generalisations, as well as generalisations that include enabling conditions, where constraints are imposed on the types of inferences that may be performed with these new types of generalisations. A new notion of attack is defined, namely alternative attack. Our approach allows arguments to be evaluated using Dung's semantics. We formally prove that instantiations of our argumentation formalism satisfy key rationality postulates [10].

The paper is structured as follows. In Section 2 we provide a conceptual analysis of reasoning about evidence. In Section 3 we present examples of analyses performed using informal reasoning tools typically used by domain experts, namely Wigmore charts and mind maps, which illustrates that both deductive and abductive inference is performed by domain experts using both causal and evidential generalisations, abstractions, and other types of generalisations. Based on these examples, in Section 4 we motivate and define our IG-formalism. In Section 5 we then define our argumentation formalism based on our IG-formalism and prove formal properties of our approach. In Section 6 we discuss related work. In Section 7 we summarise our findings and conclude.

2. Reasoning about evidence

In this section, we provide a conceptual analysis of reasoning about evidence, where we review the terminology used to describe it and introduce assumptions that demarcate the scope of the work presented in this paper. This analysis extends the analysis provided in our previous work [39] in which only causal and evidential generalisations without enablers were considered. More specifically, we now also consider abstractions and other types of generalisations, as well as generalisations that include enabling conditions. The concepts and assumptions introduced in this section are formalised in Sections 4 and 5.

Inference is the process of drawing conclusions from premises starting from the evidence, where evidence is that what has been established with certainty in the context under consideration. For instance, in the context of a legal trial, the evidence consists of that what is actually observed by a judge or jury, such as documents (e.g. police and autopsy reports) and other tangible evidence, as well as testimonial evidence [2]. Inference is often performed using domain-specific *generalisations* [2, 4, 6], also called defaults [26, 33], which capture knowledge about the world in conditional form. Generalisations can either be strict or defeasible, where defeasible generalisations are of the form *'If* $a_1, \dots, a_n$*, then usually/normally/typically b'* and strict generalisations are of the form *'If* $a_1, \dots, a_n$*, then always b'*. Here, claims $a_1, \dots, a_n$ are called the *antecedents* of the generalisation and b its *consequent*, where we assume that claims are literal propositions and that generalisations have one or more antecedents and exactly one consequent. In case a generalisation has multiple antecedents, it expresses that only the antecedents together allow us to infer the consequent. We semi-formally denote generalisations as $a_1, \dots, a_n \rightarrow b$, among others to ease the description of examples in this section and in Section 3. For defeasible generalisations, exceptional circumstances can be provided under which the generalisation may not hold, whereas strict generalisations hold without exception. An example of a (defeasible) generalisation is *'If fire, then typically smoke'*, where *'fire'* is its antecedent and *'smoke'* its consequent. An example of an exception to this generalisation is that sufficient oxygen is present for complete combustion to occur.

A distinction can be made between causal and evidential generalisations [4, 26], where instead of writing these generalisations in the form *'If ..., then ...'*, causal generalisations are written as *'*$c_1, \dots, c_n$ *usually/normally/typically cause e'* (e.g. *'fire typically causes smoke'*) and evidential generalisations are written as *'*$e_1, \dots, e_n$ *are evidence for c'* (e.g. *'smoke is evidence for fire'*). For a causal generalisation, its antecedents express causes for the consequent, and for an evidential generalisation, its consequent expresses the usual cause for its antecedents. In the context of commonsense reasoning about evidence, causal and evidential generalisations are often assumed to be defeasible (see e.g. [4, 19]); in this paper, this assumption is also made. The examples considered throughout this paper illustrate that causal and evidential generalisations are typically not strict[1].

In this paper, we also consider generalisations that are neither causal nor evidential. For instance, abstractions [4, 13] allow for reasoning at different levels of abstraction. More precisely, abstractions are of the form *'*$p_1, \dots, p_n$ *can usually/normally/typically/always be considered a specialisation of q'* (e.g. guns can usually be considered deadly weapons), where antecedents $p_1, \dots, p_n$ are considered to be more specific than the more abstract consequent q. As noted by Console and Dupré [13], abstractions are syntactically the same as causal generalisations but they are semantically different in that the antecedents of abstractions do not express causes for the consequent or vice versa. Abstractions may be defeasible (cf. [4]) but may also be strict (cf. [13]); an example of a strict abstraction is generalisation *lung_cancer* $\rightarrow_{\mathrm{a}}$ *cancer*, which states that lung cancer is a type of cancer. An example of defeasible abstraction is

[1] Note that strict generalisations such as strict rules from classical logic and definitions can be expressed using strict generalisations of type 'other' and strict abstractions.

Table 1
Table indicating for each generalisation type whether generalisations may be defeasible or strict.

	Causal generalisations	Evidential generalisations	Abstractions	Other generalisations
Defeasible	V	V	V	V
Strict	X	X	V	V

gun $\rightarrow_{\mathrm{a}}$ *deadly_weapon*, where an example of an exception to this generalisation is that the gun is a non-functional replica, or a water gun.

Another example of a different type of generalisation is a generalisation representing a mere statistical correlation, such as a correlation between homelessness and criminality. While there may be one or more confounding factors that cause both homelessness and criminality (e.g. unemployment), a domain expert may be unaware of these factors or may wish to refrain from expressing them explicitly. In this paper, we distinguish between generalisations that are causal, evidential, abstraction, or of another type, where generalisations of type 'other' may be defeasible or strict. Specifically, as this category contains all possible types of generalisations other than causal, evidential and abstraction, we allow for the option to distinguish between strict and defeasible generalisations among these generalisations. Table 1 provides an overview of the different generalisation types, where for each type it is indicated whether generalisations may be defeasible or strict. The notation $\rightarrow_{\mathrm{c}}$, $\rightarrow_{\mathrm{e}}$, $\rightarrow_{\mathrm{a}}$ and $\rightarrow_{\mathrm{o}}$ is used for the different types of generalisations, respectively.

Different types of inferences can be performed with generalisations depending on whether their antecendents or consequent are *affirmed* in that they are either observed or inferred; here, a claim is *inferred* iff it is either deductively or abductively inferred, where in deductive inference the consequent is inferred from the antecedents and in abductive inference the antecendents are inferred from the consequent. These two inference types are now considered in more detail.

2.1. Deductive inference

Inference can be performed in a deductive fashion, where from a generalisation and by affirming the antecedents, the consequent is inferred by modus ponens on the generalisation. As noted in the introduction, the term 'deduction' is used for both defeasible and strict 'forward' inference; hence, deduction is not necessarily a stronger or more reliable form of inference than abduction, which is a type of defeasible inference. Defeasible deduction can only be performed using defeasible generalisations (of any type) and not using strict generalisations (see Table 2). Strict deductive inference can only be performed using strict abstractions and strict generalisations of type 'other'. For a given instance of deductive inference, it will be explicitly specified whether it concerns strict or defeasible deductive inference.

Example 1. *Consider causal generalisation g:* fire $\rightarrow_{\mathrm{c}}$ smoke. *By affirming g's antecedent* fire*, its consequent* smoke *is defeasibly deductively inferred.* □

The following example illustrates strict deductive inference.

Example 2. *Consider strict abstraction g:* lung_cancer $\rightarrow_{\mathrm{a}}$ cancer. *Upon observing that a person has lung cancer, we can strictly deductively infer that the person has cancer using g.* □

Table 2
Table indicating for defeasible and strict generalisations of every type which types of inferences may be performed.

	Causal generalisations	Evidential generalisations	Defeasible abstractions	Strict abstractions	Defeasible other generalisations	Strict other generalisations
Defeasible deduction	V	V	V	X	V	X
Strict deduction	X	X	X	V	X	V
Abduction	V	X	V	V	X	X

Prediction [34] is a specific type of deductive inference in which the consequent of a causal generalisation is deductively inferred by affirming its antecedents. Specifically, as the antecedents of a causal generalisation express causes for the consequent, the consequent is said to be predicted from the antecedents in this case. Example 1 provides an example of prediction.

2.2. Abductive inference

Abduction [13, 18], a type of defeasible inference, can be performed using causal generalisations and abstractions: from a causal generalisation or an abstraction and by affirming the consequent, the antecedents are inferred, since if the antecedents are true it would allow us to deductively infer the consequent modus-ponens-style. Following [18], in case causes $c_1, \ldots, c_n$ and $c'_1, \ldots, c'_m$ are abductively inferred from common effect e using causal generalisations $g_1: c_1, \ldots, c_n \rightarrow_\mathrm{c} e$ and $g_2: c'_1, \ldots, c'_m \rightarrow_\mathrm{c} e$, then c_i and c'_j for $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$, $c_i \neq c'_j$ are considered to be *competing alternative explanations* for e. We assume that causes c_i (and c'_j) are not in competition among themselves.

Example 3. *Consider the following causal generalisations:*

$g_1:$ fire $\rightarrow_\mathrm{c}$ smoke*;*
$g_2:$ smoke_machine $\rightarrow_\mathrm{c}$ smoke.

*By affirming the common consequent (*smoke*),* fire *and* smoke_machine *are abductively inferred, which are then competing alternative explanations of* smoke. □

Abduction can also be performed using abstractions [4, 13], where the used abstraction can either be defeasible (cf. [4]) or strict (cf. [13]). An example of a model including strict abstractions is that of Console and Dupré [13], in which both explanatory axioms (comparable to causal generalisations) and abstraction axioms are used to explain observations. Multiple explanations that are inferred using abstraction axioms can then be considered competing alternative explanations. Note that an abductive inference step with a strict abstraction is still defeasible, as it concerns an inference step from the more abstract consequent to a more specific antecedent. Following Console and Dupré [13] and Bex [4], we allow for abduction using both strict and defeasible abstractions, where in performing abduction with abstractions $g_1: p_1, \ldots, p_n \rightarrow_\mathrm{a} q$ and $g_2: p'_1, \ldots, p'_m \rightarrow_\mathrm{a} q$ the antecedents p_i and p'_j for $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$, $c_i \neq c'_j$ are considered to be competing alternative explanations of the common consequent q. We assume that antecedents p_i (and p'_j) are not in competition among themselves.

Example 4. *Consider the following defeasible abstractions:*

$g_1:$ gun $\rightarrow_{\mathrm{a}}$ deadly_weapon*;*
$g_2:$ knife $\rightarrow_{\mathrm{a}}$ deadly_weapon.

*By affirming the common consequent (*deadly_weapon*),* gun *and* knife *are abductively inferred using generalisations g_1 and g_2, which are then competing alternative explanations of* deadly_weapon. □

The following example illustrates abductive inference with strict abstractions.

Example 5. *Consider the following strict abstractions:*

$g'_1:$ lung_cancer $\rightarrow_{\mathrm{a}}$ cancer*;*
$g'_2:$ colon_cancer $\rightarrow_{\mathrm{a}}$ cancer.

Upon observing that a person has cancer, lung_cancer *and* colon_cancer *are abductively inferred, which are then competing alternative explanations of* cancer. □

2.3. Representing causal knowledge

Abductive inference with causal generalisations and deductive inference with evidential generalisations are related: in some cases, we will accept not only causal generalisation *'c usually/normally/typically causes e'* but also evidential generalisation *'e is evidence for c'* [5, 26], which we will call the *evidential counterpart* of the causal generalisation. However, it can be argued that we only accept the evidential counterpart of a causal generalisation if c is the *usual* cause of e, where we assume that only one cause can be the usual cause of e.

Example 6. *Fire can be considered the usual cause of smoke, so we will accept both causal generalisation $g:$* fire $\rightarrow_c$ smoke *and its evidential counterpart $g':$* smoke $\rightarrow_e$ fire. *In this case, abduction with generalisation g can be encoded as deduction with generalisation g'. Because a smoke machine cannot be considered the usual cause of smoke, we will accept causal generalisation* smoke_machine $\rightarrow_c$ smoke *but we will not accept evidential generalisation* smoke $\rightarrow_e$ smoke_machine. □

Note that a causal generalisation g can only have an evidential counterpart g' in case g has a single antecedent, as we assume generalisations have a single consequent but multiple antecedents. Furthermore, as we assume that only one cause can be the usual cause of e, only one of the causal generalisations $c_1 \rightarrow_{\mathrm{c}} e$ or $c_2 \rightarrow_{\mathrm{c}} e$ can be replaced by an evidential generalisation. Hence, we do not consider c_1 and c_2 to be competing alternative explanations of e in case deductive inference is performed using evidential generalisations $e \rightarrow_{\mathrm{e}} c_1$ and $e \rightarrow_{\mathrm{e}} c_2$.

2.4. Mixed inference and inference constraints

Deductive and abductive inference can be iteratively performed, where *mixed* abductive-deductive inference is also possible.

Example 7. *Suppose that from the causal generalisation $g_1:$* fire $\rightarrow_{\mathrm{c}}$ smoke *and by affirming its consequent (*smoke*), its antecedent (*fire*) is inferred. Now, if the additional causal generalisation $g_2:$* fire $\rightarrow_{\mathrm{c}}$ heat *is provided, then its consequent (*heat*) can be deductively inferred (or predicted) as the antecedent (*fire*) has been previously abductively inferred.* □

2.4.1. Constraints on performing inference with causal and evidential generalisations

Mixed deductive inference, using both causal and evidential generalisations, can also be performed [5], but as noted by Pearl [26] care should be taken in performing mixed inference that no cause for an effect is inferred in case an alternative cause for this effect was already previously inferred.

Example 8. *Consider the example in which causal generalisation* g_1: smoke_machine $\rightarrow_c$ smoke *and evidential generalisation* g_2: smoke $\rightarrow_e$ fire *are provided. Deductively chaining these generalisations would make us infer that there is a fire when seeing a smoke machine, which is clearly undesirable.* □

Similarly, in performing mixed deductive-abductive inference, care should be taken that no cause for an effect is inferred in case an alternative cause for this effect was already previously inferred.

Example 9. *Consider Example 8, where instead of an evidential generalisation* g_2: smoke $\rightarrow_e$ fire *a causal generalisation* g_2: fire $\rightarrow_c$ smoke *is provided. Upon seeing a smoke machine, this would make us infer that there is a fire in case deductive inference and abductive inference are performed in sequence, which is again undesirable.* □

Accordingly, we wish to prohibit these types of inference patterns, and refer to the constraint that no cause for an effect should be inferred in case an alternative cause for this effect was already previously inferred as *Pearl's constraint* [26].

The above discussion can be extended to generalisations with multiple antecedents.

Example 10. *Suppose that the following generalisations are provided:*

g_1: high_body_temperature $\rightarrow_e$ fever;
g_2: smoke $\rightarrow_c$ coughing;
g_3: fever, coughing $\rightarrow_e$ pneumonia.

Upon observing that a person has high body temperature and that there is smoke, this would make us infer that the person has a fever and is coughing using generalisations g_1 *and* g_2*, respectively. In turn, this would make us infer that the person has pneumonia using generalisation* g_3*, which is undesirable: as a cause for* coughing *was already previously inferred (*smoke*), we should not be able to infer a different cause for* coughing *(*pneumonia*). Specifically,* fever *is in itself not a sufficient condition for inferring* pneumonia*:* coughing *is also necessary. Only in case a separate evidential generalisation* g_4: fever $\rightarrow_e$ pneumonia *is provided should we be able to infer* pneumonia. □

Similar problems arise in performing inference using causal generalisations with multiple antecedents. Accordingly, we wish to extend Pearl's constraint to generalisations with multiple antecedents. However, there are exceptions under which we do not wish to prohibit the aforementioned types of inference patterns, namely in case additional circumstances, also called *enabling conditions* [12], or enablers, are provided under which a causal or evidential generalisation may be used in performing inference. Generalisations that include enablers are of the general form $e_1, \ldots, e_m, a_1, \ldots, a_n \rightarrow b$, where $e_1, \ldots, e_m$ are its enablers and $a_1, \ldots, a_n$ its actual antecedents. For a causal generalisation, only its actual antecedents and not its enablers express causes for the consequent. Similarly, for an evidential generalisation its consequent only expresses the usual cause for its actual antecedents and not for its enablers. Causality is a contentious topic, and it is easy to disagree about whether an event is an actual cause or an enabler. Cheng and Novick [12] note that an event is typically viewed as an actual cause if it describes a situation that deviates from 'normal' circumstances. For instance, lighting a match is considered a cause of fire, but the presence of oxygen is typically not consider a cause of fire as it is normal that oxygen is present. This is, however, also context-dependent, and oxygen can be considered a cause of fire in situations where oxygen is typically not present (e.g. in space). We note that generalisations capture knowledge about the world as perceived by the person stating the knowledge, and that the distinction between enablers and actual causes allows domain experts to be more expressive in stating their knowledge.

The following example illustrates that deductively chaining a causal and an evidential generalisation does not lead to undesirable results for evidential generalisations that include enablers.

Example 11. *Consider the example in which the following evidential generalisation is provided:*

$g_1:$ fire, dry_wood $\rightarrow_{\mathrm{e}}$ lightning_strike.

Generalisation g_1 states that from the presence of dry_wood *and* fire *we can conclude that there may have been a lightning strike. In this case,* dry_wood *is an enabler of the generalisation, and* lightning_strike *cannot be considered a cause for antecedent* dry_wood*. Only in case fire was previously deductively inferred using a causal generalisation (e.g. $g_2:$* torch $\rightarrow_{\mathrm{c}}$ fire*) should the application of evidential generalisation g_1 be blocked. However, in case* dry_wood *was previously inferred using a causal generalisation (e.g. $g_3:$* warm_summer $\rightarrow_{\mathrm{c}}$ dry_wood*) and* fire *is not inferred using a causal generalisation, then we should be able to infer* lightning_strike *using generalisation g_1.* □

Similarly, inference may be performed using causal generalisations that include enablers, but Pearl's constraint does not need to be reconsidered in this case as illustrated by the following example.

Example 12. *Consider the example in which the following causal generalisations are provided:*

$g_1:$ torch $\rightarrow_{\mathrm{c}}$ fire;
$g_2:$ match, oxygen $\rightarrow_{\mathrm{c}}$ fire.

In this case, the presence of oxygen is an enabler of generalisation g_2, as it cannot be considered an actual cause of fire. Upon striking a match in the presence of oxygen, we can deductively infer that there is a fire using generalisation g_2. Similar to Example 9, we should now not be able to abductively infer torch *using generalisation g_1. Similarly, performing deduction and abduction in sequence using generalisations g_1 and g_2 is undesirable.* □

To summarise this section, we wish to prohibit (1) subsequent deductive inference using a causal and an evidential generalisation in case the consequent of the causal generalisation is an actual antecedent of the evidential generalisation and not an enabler, and (2) subsequent deductive and abductive inference using two causal generalisations with the same consequent. Note that, while these constraints deviate from Pearl's original constraints [26] as enabling conditions are now also taken into account, we will refer to these constraints as *Pearl's constraint* throughout this paper.

2.4.2. Constraints on performing inference with abstractions

When performing inference with abstractions, care should be taken that no version of an event at a lower level of abstraction is abductively inferred if an alternative version of this event at a lower level of abstraction was already previously inferred. In particular, performing deduction and abduction in that order with two abstractions with the same consequent leads to undesirable results.

Example 13. *Consider generalisations* $g_1\colon$ gun $\rightarrow_{\text{a}}$ deadly_weapon *and* $g_2\colon$ knife $\rightarrow_{\text{a}}$ deadly_weapon *from Example 4. Upon observing that a provided object is a* gun*, this would make us deductively infer that this object is a* deadly_weapon *using generalisation* g_1*. Upon performing abduction with* g_2*, this would make us infer that the provided object is a* knife*, which is clearly undesirable.* □

Performing abduction and deduction in that order with two abstractions with the same consequent does not lead to undesirable results.

Example 14. *Consider abstractions* $g_2\colon$ knife $\rightarrow_{\text{a}}$ deadly_weapon *and* $g_3\colon$ knife $\rightarrow_{\text{a}}$ metal_object*. Upon observing* metal_object*, we can abductively infer* knife *using generalisation* g_3*. In turn, claim* deadly_weapon *is deductively inferred using generalisation* g_2*.* □

The following example illustrates that mixed inference, using either a causal generalisation and an abstraction or an evidential generalisation and an abstraction, does not lead to undesirable results. Hence, no additional inference constraints need to be imposed.

Example 15. *Consider Example 5. Assume that in addition to strict abstractions* $g'_1\colon$ lung_cancer $\rightarrow_{\text{a}}$ cancer *and* $g'_2\colon$ colon_cancer $\rightarrow_{\text{a}}$ cancer*, causal generalisation* $g'_3\colon$ smoking $\rightarrow_{\text{c}}$ cancer *is provided. Upon observing that a person smokes, we deductively infer that the person has cancer using generalisation* g'_3*. Using generalisations* g'_1 *and* g'_2*, we can then in turn abductively infer that the person has either lung cancer or colon cancer, which are then competing alternative explanations of* cancer *(see Example 5). Note that it is not undesirable to infer* lung_cancer *or* colon_cancer *from* cancer *in this case, as* smoking *and* lung_cancer *(*colon_cancer*) are not alternative explanations of* cancer*; instead,* smoking *is a cause of* cancer*, while* lung_cancer *(*colon_cancer*) is not a cause of* cancer *but instead describes claim* cancer *at a lower level of abstraction. Similar observations can be made by replacing generalisation* g'_3 *by generalisation* $g'_4\colon$ cancer $\rightarrow_{\text{e}}$ smoking*.* □

To summarise this section, we only wish to prohibit subsequent deduction and abduction using two abstractions with the same consequent and not other inference patterns involving abstractions. Finally, note that for generalisations of type ‘other’ no additional inference constraints are imposed.

2.5. Ambiguous inference

Situations may arise in practice in which both deduction and abduction can be performed with the same causal generalisation or abstraction; the inference type is, therefore, *ambiguous*.

Example 16. *Consider generalisation* $g_1\colon$ fire $\rightarrow_{\text{c}}$ smoke*. Suppose* fire *and* smoke *are not observed but have been previously inferred, for instance via deduction using generalisations* $g_2\colon$ see_fire $\rightarrow_{\text{e}}$ fire *and* $g_3\colon$ see_smoke $\rightarrow_{\text{e}}$ smoke*, where* see_fire *and* see_smoke *are provided as evidence. Then both deduction and abduction can be performed with* g_1 *to infer* smoke *from* fire *and* fire *from* smoke*.* □

Generally, we do not wish to prohibit this type of ambiguous inference patterns as we do not consider them to be undesirable.

3. Examples of analyses performed using informal reasoning tools

In this section, we present examples of analyses performed using two tools that are typically used by domain experts, namely Wigmore charts [40] and mind maps [24, 36]. Based on these examples, we motivate and illustrate the design choices for our IG-formalism in Section 4.

3.1. *Example of an analysis performed in a Wigmore chart*

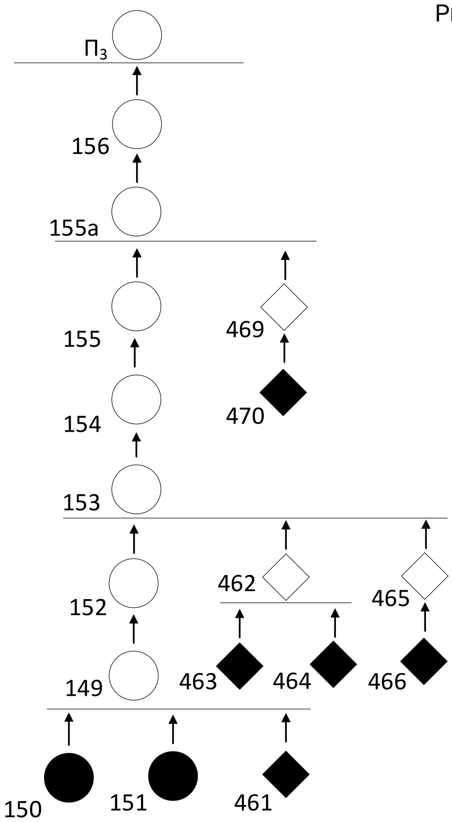
First, Wigmore charts are considered, which are diagrams familiar to legal experts in which symbols indicating hypotheses and claims are joined by lines representing relations between these hypotheses and claims. Wigmore charts were introduced by John Henry Wigmore [40] and were applied and further developed by Anderson, Schum, Twining and others (e.g. [2, 19]), who provided a modernised, more user-friendly version of Wigmore's charting method. In this section, we consider these modern versions of Wigmore charts, specifically the version adopted by Kadane and Schum [19]. In their charts, each symbol represents a unique claim. As noted by Kadane and Schum [19, p. 88], vertical arcs between nodes in the chart indicate inferences between corresponding claims, where the generalisations used in performing these inferences are not explicitly recorded in the chart. To be able to interpret whether inferences are deductive or abductive, and hence what the antecedents and consequents are of generalisations used in performing the inferences, the evidence in the chart also needs to be considered.

Example 17. *An example of a modern Wigmore chart, adapted from Kadane and Schum [19, pp. 330–331], is depicted in Figure 1, which also serves as our running example. The Wigmore chart concerns parts of an actual legal case, namely the well-known Sacco and Vanzetti case. The case concerns Sacco and Vanzetti, who were convicted for shooting and killing payroll guard Berardelli during a robbery. In this example, we only consider the part of the case concerning Sacco's consciousness of guilt. During their arrest, Sacco and Vanzetti were armed. According to the two arresting officers, Connolly and Spear, Sacco and Vanzetti made suspicious hand movements, from which the prosecution concluded that they intended to draw their concealed weapons in order to escape their arrest. This suggests that they were conscious of having committed a criminal act.*

On the right-hand side of Figure 1 the corresponding key list is depicted, which indicates for every number in the chart to which claim it corresponds. Claims provided by the defence and prosecution are represented as diamonds and circles in the chart, respectively, where nodes corresponding to the evidence are shaded. Finally, horizontal lines in the Wigmore chart indicate that information needs to be combined to draw a conclusion. □

As noted earlier, the generalisations used in performing the indicated inferences are left implicit in the chart. Instead, in their analysis of the case some of the used generalisations are indicated in the text (see e.g. [19, pp. 97–98]). For instance, generalisations used in the inferences from the testimonies are of the general form *'If a person testifying under oath tells us that event E occurred, then this event (probably, usually, often, etc,) did occur.'* [19, p. 88]. As noted by Kadane and Schum [19, pp. 74–76], in constructing their charts abduction is in some instances performed to generate interim hypotheses between the evidence and the ultimate claim Π_3. However, Kadane and Schum do not explicitly indicate which inferences in their charts are abductive and which are deductive.

Lastly, it is important to note that the manner in which claims and links conflict is not precisely specified in Kadane and Schum's Wigmore charts, as also observed by Bex and colleagues [6] in formalising



PROSECUTION
149. Following his arrest, Sacco attempted to put his hand under his overcoat.
150. Connolly's testimony to 149.
151. Spear's testimony to 149.
152. Sacco intended to draw his concealed weapon.
153. Sacco intended to use his weapon on the arresting officers.
154. Sacco intended to escape from his arrest.
155. Sacco was conscious of having committed a criminal act.
155a. Sacco was conscious of having been involved in a robbery and shooting.
156. Sacco was conscious of having been involved in the robbery and shooting that took place in South Braintree.
Π_3. It was Sacco who, with the assistence of Vanzetti, intentionally fired shots that took the life of Berardelli during the robbery and shooting that took place in South Braintree.

DEFENCE
461. Sacco's testimony to denying 149.
462. Sacco carried a weapon because he intended to shoot rabbits with it.
463. Sacco's testimony to 462.
464. Sacco's wife's testimony to 462.
465. Sacco carried a weapon because of his duties as a night watchman.
466. Sacco's testimony to 465.
469. Sacco believed he was being arrested because of his political beliefs.
470. Sacco's testimony to 469.

Fig. 1. Wigmore chart concerning Sacco's consciousness of guilt, along with the corresponding key list, adapted from Kadane and Schum [19, pp. 330–331].

such Wigmore charts as Pollock-style arguments [28]. For instance, multiple interpretations of the conflicts between the defence's claims 462 and 465 and the prosecution's claims 152 and 153 are possible. One possible interpretation is that 462 and 465 indicate support for the negation of claim 153: as Sacco carried his weapon for an innocent reason (either 462 or 465), he intended to surrender his weapon and, therefore, did not intend to use it. Alternatively, 462 and 465 can be considered competing alternative explanations of 152, and hence be interpreted as exceptions to the performed inference step from 152 to 153. Specifically, as Sacco carried his weapon for an innocent reason (462 or 465), this caused him to draw his weapon (152) with the intention of surrendering it.

3.2. Example of an analysis performed using a mind mapping tool

Next, we present an example of an analysis performed using a mind mapping tool [24], which is an example of a tool typically used by domain experts, for instance in crime analysis [36]. A mind map usually takes the shape of a diagram in which hypotheses and claims are represented by boxes and underlined text, and undirected edges symbolise relations between these hypotheses and claims. An example is depicted in Figure 2, which is based on a standard template used by the Dutch police for criminal cases involving the suspicious death of a person. The mind map represents various scenario-elements and the crime analyst uses evidence to support or oppose these elements, indicated in the mind

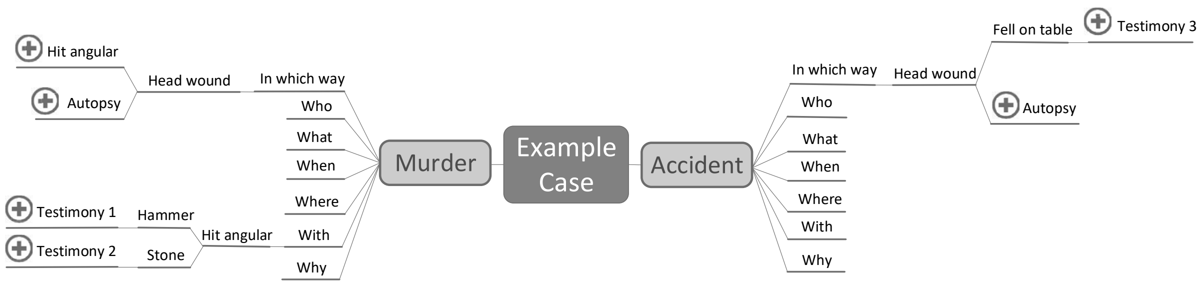


Fig. 2. Example of a partially filled in mind map.

map by plus and minus symbols, respectively. Compared to Wigmore charts, which offer a wide range of symbols and arcs to allow users to be expressive and precise in modelling legal reasoning, mind maps are less precise and are used to obtain an overview of different possible alternative scenarios. In the following example, only supporting evidence is considered, which allows us to focus on the manner in which competing alternative explanations are captured in mind maps.

Example 18. *An example of a partially filled in mind map is depicted in Figure 2. In this example case, a body was found; we are interested in the cause of death of this person. First, high-level hypothesis 'Murder' is examined. According to witness testimony (*Testimony 1*), the person was hit with a hammer (*Hammer*); however, according to another testimony (*Testimony 2*), the person was hit with a stone (*Stone*). By means of plus symbols and undirected edges connecting the evidence to these claims, it is indicated that claims* Testimony 1 *and* Testimony 2 *support claims* Hammer *and* Stone*, respectively.* Hammer *and* Stone *are connected via undirected edges to* Hit angular*, which indicates that hammers and stones can generally be considered to be angular. In turn, claim* Hit angular *is connected to the 'With' question to indicate that it provides an answer to this question. As an answer to the 'In which way' question, it is indicated that the person died because of a head wound (*Head wound*), which is again supported by the claim that the person was hit with an angular object (*Hit angular*). An autopsy report (*Autopsy*) further supports claim* Head wound*.*

Next, high-level hypothesis 'Accident' is examined, which provides a competing alternative explanation of Head wound*. As an answer to the 'In which way' question, it is again indicated that the person died because of a head wound and that this claim is supported by* Autopsy*; however, in contrast to the answer to this question for high-level hypothesis 'Murder', it is indicated that the head wound was caused as the person fell on a table by accident (*Fell on table*), a claim supported by* Testimony 3. □

As the edges in a mind map are undirected, it is unclear from the graphical representation alone which types of generalisations and inferences were used in constructing this map. Establishing this with certainty would require directly consulting the domain experts involved in constructing the chart. We note, however, that the reasoning performed in constructing this mind map can be interpreted in at least two possible ways. One interpretation is that the domain expert first (preliminarily) inferred that the person died because of a head wound from the autopsy report via deduction using the evidential generalisation g_1: *Autopsy* $\rightarrow_{\text{e}}$ *Head wound*, and then abductively inferred *Hit angular* using the causal generalisation g_2: *Hit angular* $\rightarrow_{\text{c}}$ *Head wound*. In turn, *Hammer* and *Stone* are abductively inferred from *Hit angular* using the abstractions g_3: *Hammer* $\rightarrow_{\text{a}}$ *Hit angular* and g_4: *Stone* $\rightarrow_{\text{a}}$ *Hit angular*. These two claims are then competing alternative explanations of *Hit angular* and are subsequently grounded in evidence, namely via deduction from the testimonies using evidential generalisations $g_5:$ *Testimony 1* $\rightarrow_{\mathrm{e}}$ *Hammer* and $g_6:$ *Testimony 2* $\rightarrow_{\mathrm{e}}$ *Stone*. An alternative interpretation is that the mind map was constructed iteratively from the evidence, where from the testimonies the claims *Hammer* and *Stone* are inferred via deduction using generalisations g_5 and g_6. Claim *Hit angular* is then inferred modus-ponens style: from abstractions g_3 and g_4 and the previously inferred antecedents, the consequent is deductively inferred. In this way, *Hammer* and *Stone* are not in competition for *Hit angular*.

This example illustrates that the types of generalisations and inferences involved in the analysis of a case using a mind mapping tool are typically left implicit. Similarly, the manner in which claims and links conflict is not precisely specified in mind maps: in particular, conflicts between competing alternative explanations are not explicitly indicated in the graph.

4. The information graph formalism

The examples from Section 3 make it plausible that both deductive and abductive inference is performed by domain experts when performing analyses using reasoning tools they are familiar with. In performing such analyses, the used generalisation, as well as the inference type (deduction, abduction), are left implicit. Furthermore, the assumptions of domain experts underlying their analyses are typically not explicitly stated, making these analyses ambiguous to interpret. For current purposes, we wish to provide a precise account of the interplay between the different types of inferences and generalisations that formalises and disambiguates these analyses in a manner that makes the used generalisations explicit. *Information graphs* (IGs), which we define in Section 4.1, are knowledge representations that explicitly describe generalisations in the graph. In constructing an IG from an analysis performed using a tool, an interpretation step may be required; we provide examples of this interpretation step by discussing possible formalisations of the Wigmore chart of Section 3.1 and the mind map of Section 3.2. In Section 4.2, we define how deductive and abductive inferences can be read from IGs given the evidence, based on our conceptual analysis of reasoning about evidence (Section 2). Compared to our previously proposed IG-formalism [39] in which only causal and evidential generalisations were considered, abstractions and other types of generalisations are now also considered, as well as generalisations that include enabling conditions, where constraints are imposed on the types of inferences that may be performed with these new types of generalisations.

4.1. Information graphs

First, the syntax of IGs is defined. Throughout this paper, boldface is used to indicate sets used in the IG-formalism.

Definition 1 (Information graph)**.** *An* information graph *(IG) is a directed graph* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$*, where* $\mathbf{P}$ *is a set of nodes representing propositions from a propositional language consisting of only literals and that is closed under classical negation, where the negation symbol is denoted by* $\neg$*.* $\mathbf{A}$ *is a set of (hyper)arcs that divides into three pairwise disjoint subsets* $\mathbf{G}$*,* $\mathbf{N}$ *and* $\mathbf{X}$ *of generalisation arcs, negation arcs and exception arcs, defined in Definitions 2, 4, and 5, respectively.*

For IGs, there is a one-to-one correspondence between nodes and propositions, generalisation arcs and generalisations, exception arcs and exceptions, and negation arcs and negations. Throughout this paper, in the context of IGs, the terms 'node' and 'proposition', 'generalisation arc' and 'generalisation', 'exception arc' and 'exception', and 'negation arc' and 'negation' are therefore used interchangeably. We write $p = -q$ in case $p = \neg q$ or $q = \neg p$. Finally, note that while we currently only consider classical negation, our IG-formalism may be extended in future work to allow for more general notions of conflicts such as contrariness (cf. [21]).

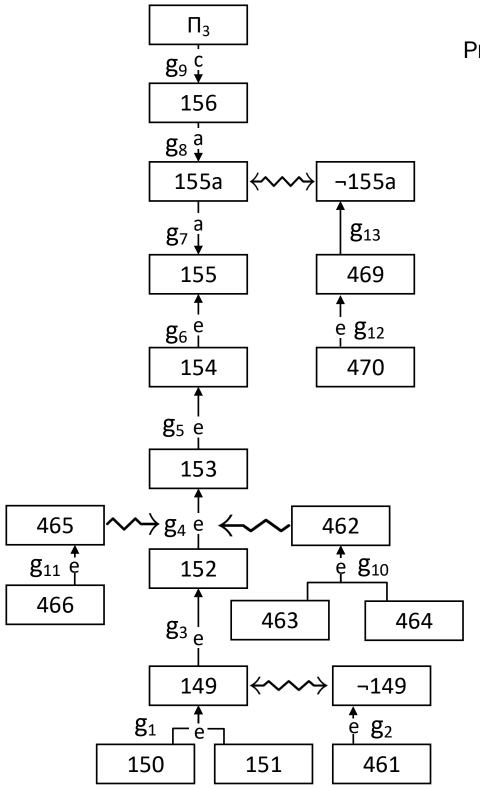
Definition 2 (Generalisation arc)**.** *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG. A* generalisation arc *$g \in \mathbf{G} \subseteq \mathbf{A}$ is a directed (hyper)arc $g\colon \{p_1, \ldots, p_n\} \to p$, indicating a generalisation with antecedents $\mathbf{P_1} = \{p_1, \ldots, p_n\} \subseteq \mathbf{P}$ and consequent $p \in \mathbf{P} \setminus \mathbf{P_1}$. Here, propositions in $\mathbf{P_1}$ are called the* tails *of g, denoted by $\mathbf{Tails}(g)$, and p is called the* head *of g, denoted by $Head(g)$. $\mathbf{G}$ divides into four pairwise disjoint subsets $\mathbf{G}^{\mathrm{c}}$, $\mathbf{G}^{\mathrm{e}}$, $\mathbf{G}^{\mathrm{a}}$ and $\mathbf{G}^{\mathrm{o}}$ of causal generalisation arcs, evidential generalisation arcs, abstraction arcs, and all other types of generalisation arcs, respectively. Generalisations in $\mathbf{G}^{\mathrm{c}}$ and $\mathbf{G}^{\mathrm{e}}$ are defeasible, $\mathbf{G}^{\mathrm{a}}$ divides into disjoint subsets $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}}$ and $\mathbf{G}^{\mathrm{a}}_{\mathrm{d}}$ of strict and defeasible abstraction arcs, respectively, and $\mathbf{G}^{\mathrm{o}}$ divides into disjoint subsets $\mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ and $\mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$ of strict and defeasible other types of generalisation arcs, respectively. For $g \in \mathbf{G}$, $\mathbf{Tails}(g)$ divides into disjoint subsets $\mathbf{Enabler}(g)$ and $\mathbf{Ant}(g)$ of propositions representing enabling conditions and actual antecedents of the generalisation, respectively, where for $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}}$ it holds that $\mathbf{Ant}(g) \neq \emptyset$ and possibly $\mathbf{Enabler}(g) = \emptyset$, and for $g \in \mathbf{G}^{\mathrm{a}} \cup \mathbf{G}^{\mathrm{o}}$ it holds that $\mathbf{Enabler}(g) = \emptyset$ (i.e. $\mathbf{Tails}(g) = \mathbf{Ant}(g)$).*

Curly brackets are omitted in case $|\mathbf{Tails}(g)| = 1$. In figures in this paper, generalisation arcs are denoted by solid (hyper)arcs, which are labelled 'c' for $g \in \mathbf{G}^{\mathrm{c}}$, 'e' for $g \in \mathbf{G}^{\mathrm{e}}$, and 'a' for $g \in \mathbf{G}^{\mathrm{a}}$, where 'o' labels for $g \in \mathbf{G}^{\mathrm{o}}$ are omitted.

In accordance with our assumptions stated in Section 2, causal and evidential generalisations are defeasible and can include enablers. Abstractions and other types of generalisations can either be strict or defeasible. A causal generalisation $g\colon c \to e$ may have an evidential counterpart of the form $g'\colon e \to c$ (see Section 2.3), but only if c is the usual cause of e. Definition 2 does not prohibit the coexistence of a causal generalisation $g\colon c \to e$ and its evidential counterpart $g'\colon e \to c$ in an IG, and inferences can be read from IGs including both generalisations without yielding anomalous results; hence, both generalisations may be included if considered desirable. However, it should be noted that g and g' represent the same knowledge, and that care should be taken in for instance modelling exceptions to generalisations (see Definition 5), as an exception to g can also be considered an exception to g'. Ultimately, it is the responsibility of the knowledge engineer in consultation with the domain expert to decide which knowledge to include in the IG and to ensure this knowledge is correctly and consistently represented.

In the following example, the Wigmore chart of Section 3.1 is modelled as an IG.

Example 19. *In Figure 3, an IG is depicted for a possible interpretation of the Wigmore chart of Figure 1. This interpretation is based on a previous interpretation of this Wigmore chart as a preliminary version of an IG in which only causal and evidential information is considered and the roles of generalisation and inference are not separated [37]. For every claim p in the Wigmore chart, a proposition node p is included in $\mathbf{P}$. As noted by Kadane and Schum [19, p. 88], the generalisations used in the inferences from the testimonies are evidential. As propositions $150, 151, 463, 464, 466$ and 470 denote testimonies, the IG includes generalisation arcs $g_1\colon \{150, 151\} \to 149$; $g_{10}\colon \{463, 464\} \to 462$, $g_{11}\colon 466 \to 465$ and $g_{12}\colon 470 \to 469$ in $\mathbf{G}^{\mathrm{e}}$. Here, testimonies 150, 151 and 463, 464 are combined in the antecedents of generalisations g_1 and g_{10}, respectively, as these sets of propositions concern testimonies to the same claim. As 461 concerns Sacco's testimony to denying 149, proposition $\neg 149$ is included in $\mathbf{P}$ and generalisation arc $g_2\colon 461 \to \neg 149$ is included in $\mathbf{G}^{\mathrm{e}}$.*



PROSECUTION
149. Following his arrest, Sacco attempted to put his hand under his overcoat.
150. Connolly's testimony to 149.
151. Spear's testimony to 149.
152. Sacco intended to draw his concealed weapon.
153. Sacco intended to use his weapon on the arresting officers.
154. Sacco intended to escape from his arrest.
155. Sacco was conscious of having committed a criminal act.
155a. Sacco was conscious of having been involved in a robbery and shooting.
156. Sacco was conscious of having been involved in the robbery and shooting that took place in South Braintree.
Π_3. It was Sacco who, with the assistence of Vanzetti, intentionally fired shots that took the life of Berardelli during the robbery and shooting that took place in South Braintree.

DEFENCE
461. Sacco's testimony to denying 149.
462. Sacco carried a weapon because he intended to shoot rabbits with it.
463. Sacco's testimony to 462.
464. Sacco's wife's testimony to 462.
465. Sacco carried a weapon because of his duties as a night watchman.
466. Sacco's testimony to 465.
469. Sacco believed he was being arrested because of his political beliefs.
470. Sacco's testimony to 469.

Fig. 3. IG corresponding to an interpretation of the Wigmore chart of Figure 1, where 'e' labels denote evidential generalisations, 'c' labels denote causal generalisations, 'a' labels denote abstractions, ⇜⇝ is a negation arc and ⇝ is an exception arc.

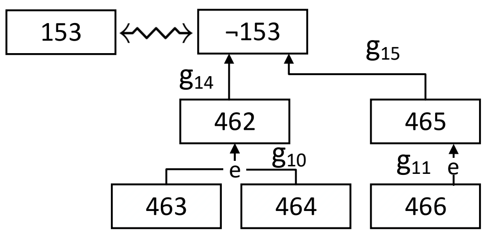


Fig. 4. Adjustment to part of the IG of Figure 3, where 462 and 465 indicate support for ¬153.

Kadane and Schum do not indicate which (types of) generalisations were used in performing the inferences between propositions 149 *and* Π_3*. We note that the inferences between* 149 *and* 155 *fit a so-called episode scheme for intentional actions [4, p. 64], a story scheme in which someone's psychological state causes them to form certain goals, which in turn lead to actions that have consequences. In this case, Sacco intended to escape from his arrest (*154*; goal) as he was conscious of having committed a criminal act (*155*; psychological state); therefore, we consider* 155 *a cause of* 154*. Sacco's intention to use his weapon (*153*) can then be considered a sub-goal of* 154 *and his intention to draw his concealed weapon (*152*) a further sub-goal of* 153*. Sacco's intention to draw his weapon (*152*) caused Sacco to attempt to put his hand under his overcoat (*149*; action); therefore, we consider* 152 *a cause of* 149*. The IG there-*

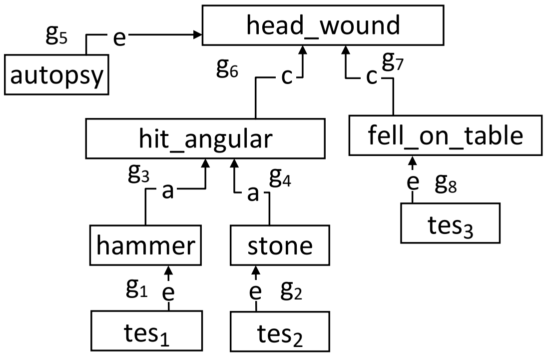


Fig. 5. IG corresponding to a possible interpretation of the mind map of Figure 2.

fore includes generalisation arcs $g_3\colon 149 \to 152$*;* $g_4\colon 152 \to 153$*;* $g_5\colon 153 \to 154$ *and* $g_6\colon 154 \to 155$ *in* $\mathbf{G}^{\mathrm{e}}$ *to denote these generalisations.*

Proposition 155 *can be considered an abstraction of* 155*a: being involved in a robbery and shooting can generally be considered committing a criminal act. The involved generalisation is defeasible: involvement in a robbery and shooting does not imply that this involvement is of a criminal nature, as it may also imply that the person under consideration is the victim. Proposition* 155*a can be considered a strict abstraction of* 156*, as at a higher level of abstraction being conscious of having been involved in the specific robbery and shooting that took place in South Braintree can be considered being conscious of having been involved in a robbery and shooting.* Π_3 *can be considered a cause of* 156*: committing a specific robbery and shooting typically causes a person (in this case Sacco) to be conscious of having been involved in this act. Therefore, generalisation* $g_7\colon 155a \to 155$ *is included in* $\mathbf{G}^{\mathrm{a}}_{\mathrm{d}}$*,* $g_8\colon 156 \to 155a$ *in* $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}}$*, and* $g_9\colon \Pi_3 \to 156$ *in* $\mathbf{G}^{\mathrm{c}}$*. Finally, from* 469 *(Sacco believed he was being arrested because of his political beliefs), we can conclude that Sacco was not conscious of having been involved in a robbery and shooting (*$\neg 155a$*). We consider the relation between* 469 *and* $\neg 155a$ *to be defeasible and neither causal nor evidential nor an abstraction, and therefore include* $g_{13}\colon 469 \to \neg 155a$ *in* $\mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$*.* □

In the following example, the mind map of Section 3.2 is modelled as an IG.

Example 20. *Consider Figure 5, which depicts an IG for a possible interpretation of the mind map of Figure 2. The generalisations used in the inferences from the testimonies, as well as from* autopsy*, are considered to be evidential; therefore, generalisation arcs* g_1, g_2, g_5 *and* g_8 *are included in* $\mathbf{G}^{\mathrm{e}}$*. The relation between* hammer *(*stone*) and* hit_angular *is neither causal nor evidential; instead, generalisation arcs* g_3 *and* g_4 *are included in* $\mathbf{G}^{\mathrm{a}}_{\mathrm{d}}$ *to express that, at a higher level of abstraction, both hammers and stones can generally be considered angular objects. These generalisations are defeasible as not all hammers and stones are angular. Finally,* hit_angular *and* fell_on_table *can both be considered causes of* head_wound*; therefore, generalisation arcs* g_6 *and* g_7 *are included in* $\mathbf{G}^{\mathrm{c}}$*.* □

The following example illustrates generalisation arcs that include enabling conditions.

Example 21. *Consider* $g'_7\colon$ {fell_on_table, no_helmet} $\to$ head_wound *in* $\mathbf{G}^{\mathrm{c}}$*, which is an adjustment to generalisation* g_7 *of Example 20 which states that falling on a table causes a head wound in case you are not wearing a helmet. As in Example 20, proposition* fell_on_table *expresses a cause for* head_wound *and hence,* fell_on_table *is included in* $\mathbf{Ant}(g'_7)$*. Proposition* no_helmet *does not express*

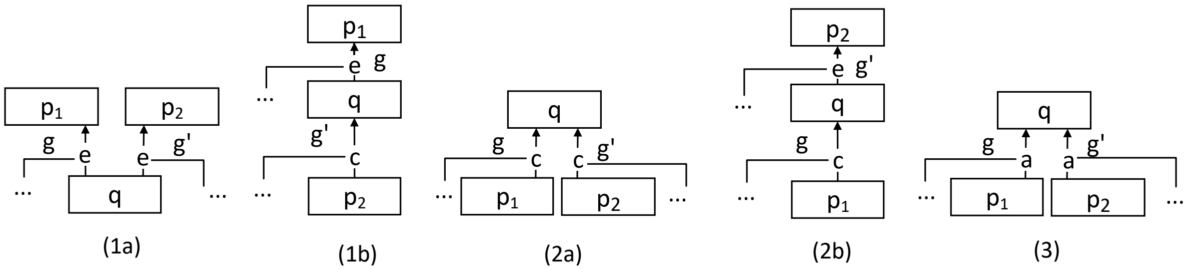


Fig. 6. Illustration of the terminology used in Definition 3.

a cause for head_wound *and can thus be considered an enabler of* g'_7*; therefore,* no_helmet *is included in* $\mathbf{Enabler}(g'_7)$*. It should be noted that, while* no_helmet *does not express a cause for the consequent, it still is a necessary condition of generalisation* g'_7*.* □

Specific configurations of generalisations express that two propositions are *alternative explanations* of a common proposition, as captured by Definition 3. The terminology used is illustrated in Figure 6.

Definition 3 (Alternative explanations)**.** *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG. Then* $p_1, p_2 \in \mathbf{P}$ *are* alternative explanations *of* $q \in \mathbf{P}$*, as indicated by generalisations* g *and* g' *in* $\mathbf{G}$*, iff one of the following holds:*

(1) $g \in \mathbf{G}^{\mathrm{e}}$, $Head(g) = p_1$, $q \in \mathbf{Ant}(g)$*, and either:*
  1a) $g' \in \mathbf{G}^{\mathrm{e}}$, $g' \neq g$, $Head(g') = p_2$, $q \in \mathbf{Ant}(g')$*, or;*
  1b) $g' \in \mathbf{G}^{\mathrm{c}}$, $Head(g') = q$, $p_2 \in \mathbf{Ant}(g')$.
(2) $g \in \mathbf{G}^{\mathrm{c}}$, $Head(g) = q$, $p_1 \in \mathbf{Ant}(g)$*, and either:*
  2a) $g' \in \mathbf{G}^{\mathrm{c}}$, $g' \neq g$, $Head(g') = q$, $p_2 \in \mathbf{Ant}(g')$*, or;*
  2b) $g' \in \mathbf{G}^{\mathrm{e}}$, $Head(g') = p_2$, $q \in \mathbf{Ant}(g')$.
(3) $g \in \mathbf{G}^{\mathrm{a}}$, $Head(g) = q$, $p_1 \in \mathbf{Tails}(g)$ *and* $g' \in \mathbf{G}^{\mathrm{a}}$, $g' \neq g$, $Head(g') = q$, $p_2 \in \mathbf{Tails}(g')$.

Note that cases 1b and 2b are symmetrical in terms of p_1 and p_2 and the used generalisations; we opt to keep the distinction between these two cases as they simplify the proof of Proposition 1. In case 1a, q is an actual antecedent and not an enabler of both $g \in \mathbf{G}^{\mathrm{e}}$ and $g' \in \mathbf{G}^{\mathrm{e}}$; hence, both p_1 and p_2 are actual causes of q. Assuming that g and g' both have multiple actual antecedents in case 1a, then p_1 and p_2 are alternative explanations of every proposition $q \in \mathbf{Ant}(g) \cap \mathbf{Ant}(g')$. Hence, it is meaningful to define alternative explanations in the context in which generalisations have non-singleton sets of actual antecedents; this similarly holds for the other cases of Definition 3. In case 1b, p_2 is an actual antecedent and not an enabler of $g' \in \mathbf{G}^{\mathrm{c}}$ and thus a cause of q, and q is an actual antecedent and not an enabler of $g \in \mathbf{G}^{\mathrm{e}}$ and thus p_1 is a cause of q. In case 2a, p_1 and p_2 are actual antecedents of $g \in \mathbf{G}^{\mathrm{c}}$ and $g' \in \mathbf{G}^{\mathrm{c}}$, respectively; hence, both p_1 and p_2 are actual causes of q. Finally, in case 3, p_1 and p_2 are antecedents of $g \in \mathbf{G}^{\mathrm{a}}$ and $g' \in \mathbf{G}^{\mathrm{a}}$ with the same consequent q, and hence are alternative explanations of q.

Example 22. *Consider the IG of Figure 5. According to condition* 2a *of Definition 3,* hit_angular *and* fell_on_table *are alternative explanations of* head_wound *as indicated by generalisations* g_6 *and* g_7*. Similarly, according to condition* 3 *of Definition 3,* hammer *and* stone *are alternative explanations of* hit_angular *as indicated by generalisations* g_3 *and* g_4*.* □

A negation arc captures a conflict between a proposition and its negation expressed in an IG.

Definition 4 (Negation arc)**.** *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG. A* negation arc $n \in \mathbf{N} \subseteq \mathbf{A}$ *is a bidirectional arc* $n\colon p \leftrightsquigarrow q$ *in* $G_{\mathcal{I}}$ *that exists between a pair* $p, q \in \mathbf{P}$ *iff* $q = -p$.

Example 23. *Consider the running example. As both* 149 *and* ¬149 *are included in the IG of Figure 3, negation arc* $n_1\colon 149 \leftrightsquigarrow \neg 149$ *is also included in the graph. Similarly, the IG of Figure 3 includes negation arc* $n_2\colon 155a \leftrightsquigarrow \neg 155a$. *As noted in Section 3.1, one possible interpretation of the conflicts between propositions* 462, 465 *and* 153 *is that* 462 *and* 465 *indicate support for* ¬153. *Accordingly, generalisations* $g_{14}\colon 462 \rightarrow \neg 153$ *and* $g_{15}\colon 465 \rightarrow \neg 153$ *can be included, as depicted in the adjusted IG of Figure 4. As these generalisations are defeasible and neither causal nor evidential nor an abstraction,* g_{14} *and* g_{15} *are included in* $\mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$. *Negation arc* $n_3\colon 153 \leftrightsquigarrow \neg 153$ *is then included in* $\mathbf{N}$. *An alternative interpretation of these conflicts is provided in Example 24.* □

As defeasible generalisations do not hold universally, exceptional circumstances can be provided under which such a generalisation may not hold; hence, we allow exceptions to defeasible generalisations to be specified in IGs by means of exception arcs.

Definition 5 (Exception arc)**.** *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG. An* exception arc $x \in \mathbf{X} \subseteq \mathbf{A}$ *is a hyperarc* $x\colon p \rightsquigarrow g$, *where* $p \in \mathbf{P}$ *is called an* exception *to defeasible generalisation* $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$.

An exception arc directed from p to g indicates that p provides exceptional circumstances under which g may not hold.

Example 24. *Consider the running example. Instead of interpreting the conflicts between propositions* 462, 465 *and* 153 *as negations (see Example 23), an alternative interpretation is that* 462 *and* 465 *indicate exceptions to generalisation* $g_4 \in \mathbf{G}^{\mathrm{e}}$. *Specifically,* 462 *and* 465 *can be considered competing alternative explanations of* 152*: as Sacco carried his weapon for an innocent reason (*462 *or* 465*), this caused him to draw his weapon (*152*) with the intention of surrendering it. In Figure 3, these exceptions are indicated by curved hyperarcs* $x_1\colon 462 \rightsquigarrow g_4$ *and* $x_2\colon 465 \rightsquigarrow g_4$ *in* $\mathbf{X}$. □

4.2. Reading inferences from information graphs

We now define how deductive and abductive inferences can be performed with constructed IGs. By itself, a generalisation arc only expresses that the tails together allow us to infer the head in case this generalisation is used in deductive inference, or that the tails together can be inferred from the head in case of abductive inference. Only when considering the available evidence can directionality of inference actually be read from the graph.

Definition 6 (Evidence set)**.** *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG. An* evidence set *is a subset* $\mathbf{E} \subseteq \mathbf{P}$ *for which it holds that for every* $p \in \mathbf{E}$, $\neg p \notin \mathbf{E}$.

The restriction that for every $p \in \mathbf{E}$ it holds that $\neg p \notin \mathbf{E}$ ensures that not both a proposition and its negation are observed.

In figures in this paper, nodes in $G_{\mathcal{I}}$ corresponding to elements of $\mathbf{E}$ are shaded and all shaded nodes correspond to elements of $\mathbf{E}$. We emphasise that various evidence sets $\mathbf{E}$ can be used to establish (different) inferences from the same IG.

Example 25. *In Figure 1, the evidence consists of the testimonies. In Figures 7 and 8, the IGs of Figures 3 and 4 are again depicted with nodes in* $\mathbf{E} = \{150, 151, 461, 463, 464, 466, 470\}$ *shaded.* □

We now define when we consider configurations of generalisation arcs and evidence to express deductive and abductive inference.

4.2.1. Deductive inference

First, we specify under which conditions we consider a configuration of generalisation arcs and evidence to express deductive inference, where strict and defeasible deduction are distinguished.

Definition 7 (Deductive inference). *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG, and let* $\mathbf{E} \subseteq \mathbf{P}$ *be an evidence set. Let* $p_1, \ldots, p_n, q \in \mathbf{P}$*, with* $q \notin \mathbf{E}$*. Then given* $\mathbf{E}$*,* q *is* deductively inferred from propositions $p_1, \ldots, p_n$ using a generalisation $g\colon \{p_1, \ldots, p_n\} \rightarrow q$ in $\mathbf{G}$ *iff* $\forall p_i$*,* $i = 1, \ldots, n$*:*

(1) $p_i \in \mathbf{E}$*, or;*
(2) p_i *is deductively inferred from propositions* $r_1, \ldots, r_m \in \mathbf{P}$ *using a generalisation* $g'\colon \{r_1, \ldots, r_m\} \rightarrow p_i$*, where* $g' \notin \mathbf{G}^{\mathrm{c}}$ *if* $g \in \mathbf{G}^{\mathrm{e}}$*,* $p_i \notin \mathbf{Enabler}(g)$*, or;*
(3) p_i *is abductively inferred from a proposition* $r \in \mathbf{P}$ *using a generalisation* $g'\colon \{p_i, r_1, \ldots, r_m\} \rightarrow r$ *in* $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$*,* $g \neq g'$*,* $r_1, \ldots, r_m \in \mathbf{P}$ *(see Definition 8).*

Here, proposition q *is* defeasibly deductively inferred from $p_1, \ldots, p_n$*, denoted* $p_1, \ldots, p_n \twoheadrightarrow_g q$*, iff* $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$*, and proposition* q *is* strictly deductively inferred from $p_1, \ldots, p_n$*, denoted* $p_1, \ldots, p_n \rightharpoonup_g q$*, iff* $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$.

For ease of reference, symbols $\twoheadrightarrow$ and $\rightharpoonup$ are annotated with the name of the generalisation used in performing a defeasible or strict inference. In accordance with our assumptions stated in Section 2.1, deduction can be performed using all types of generalisations in $\mathbf{G}$, where strict deduction can only be performed using strict abstractions and strict other types of generalisations. The condition $q \notin \mathbf{E}$ ensures that deduction cannot be performed with a generalisation to infer its consequent in case its consequent is already observed. Deduction can only be performed using a generalisation $g \in \mathbf{G}$ to infer its consequent $Head(g)$ from its antecedents $\mathbf{Tails}(g)$ in case every antecedent $p_i \in \mathbf{Tails}(g)$ has been affirmed in that either p_i is observed (i.e. $p_i \in \mathbf{E}$), p_i itself is deductively inferred, or p_i is abductively inferred. In correspondence with Pearl's constraint (see Section 2.4.1), we assume in condition 2 that a proposition $q \in \mathbf{P}$ cannot be deductively inferred from $p_1, \ldots, p_n \in \mathbf{P}$ using a generalisation $g \in \mathbf{G}^{\mathrm{e}}$ if at least one of its actual antecedents $p_i \in \mathbf{Ant}(g)$ is deductively inferred using a generalisation $g' \in \mathbf{G}^{\mathrm{c}}$. In this case, q and propositions $r_i \in \mathbf{Ant}(g')$ are considered alternative explanations of p_i as indicated by g and g' (Definition 3, case 1b or case 2b). Condition 3 of Definition 7 is explained in Section 4.2.3, after abductive inference is defined.

Example 26. *In the IG of Figure 7, given* $\mathbf{E}$ *propositions* 149, $\neg$149, 462, 465 *and* 469 *are defeasibly deductively inferred from* 150 *and* 151, 461, 463 *and* 464, 466*, and* 470 *using generalisations* g_1*,* g_2*,* g_{10}*,* g_{11}*, and* g_{12}*, respectively, as* $150, 151, 461, 463, 464, 466, 470 \in \mathbf{E}$ *(Definition 7, condition* 1*). Proposition* 152 *is then defeasibly deductively inferred from* 149 *using* g_3*, as* 149 *is deductively inferred (Definition 7, condition* 2*). Propositions* 153, 154 *and* 155 *are then iteratively defeasibly deductively inferred using generalisations* g_4, g_5 *and* g_6*, respectively. Finally, from* 469, $\neg$155*a is defeasibly deductively inferred using* g_{13}*, as* 469 *is deductively inferred.* □

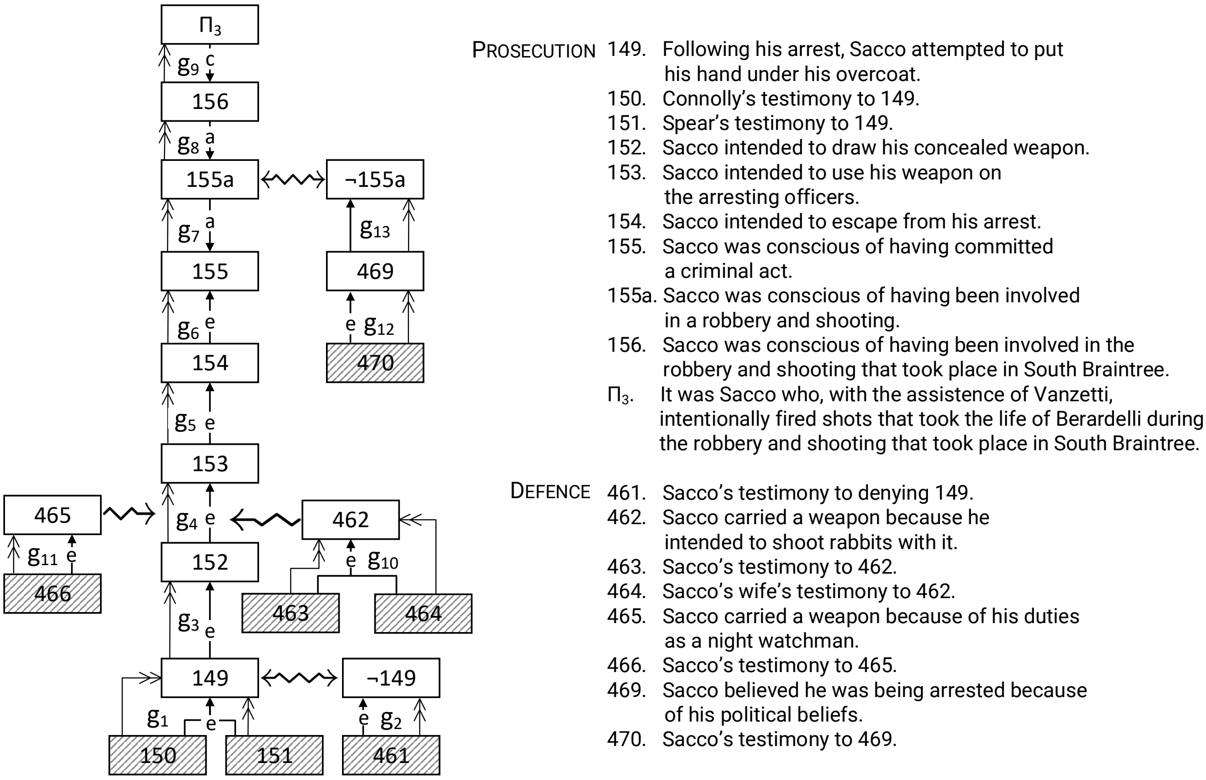


Fig. 7. The IG of Figure 3, where evidence set $\mathbf{E}$ (shaded) and resulting inference steps ($\twoheadrightarrow$) are also indicated.

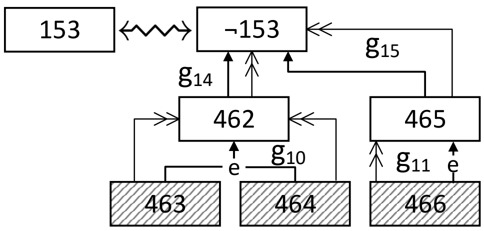


Fig. 8. The IG of Figure 4, where evidence set $\mathbf{E}$ (shaded) and resulting inference steps ($\twoheadrightarrow$) are also indicated.

The following example illustrates strict deductive inference.

Example 27. *Consider Example 2 from Section 2.1. In this example, generalisation arc* g: lung_cancer $\rightarrow$ cancer *is included in* $\mathbf{G}^{\mathrm{a}}_{\mathrm{S}}$. *As* lung_cancer $\in \mathbf{E}$, cancer *is strictly deductively inferred from* lung_cancer *(Definition 7, condition* 1*).* □

The next example illustrates the restrictions put on performing deduction in our IG-formalism.

Example 28. *Figure 9a depicts an example of an IG in which* q *cannot be deductively inferred from* p *using* g_1*, as* $Head(g_1) = q \in \mathbf{E}$*. In Figure 9b,* q *cannot be deductively inferred from* p_1 *and* p_2 *using* g_1*, as* $p_2 \notin \mathbf{E}$ *and* p_2 *is neither deductively nor abductively inferred.*

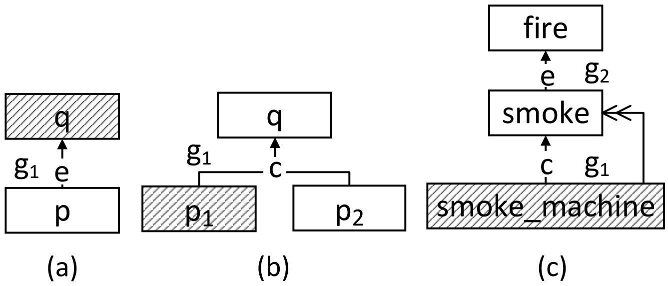


Fig. 9. Examples of IGs illustrating the restrictions put on performing deduction within our IG-formalism (a-c).

In Figure 9c, Example 8 illustrating Pearl's constraint is modelled. As smoke_machine $\in \mathbf{E}$, smoke *is deductively inferred from* smoke_machine *using* g_1 *by condition* 1 *of Definition 7.* fire *cannot in turn be inferred from* smoke *using* g_2 *by condition* 2 *of Definition 7, as* $g_2 \in \mathbf{G}^{\mathrm{e}}$ *and* smoke *is deductively inferred using* $g_1 \in \mathbf{G}^{\mathrm{c}}$. □

4.2.2. Abductive inference

Next, we specify under which conditions we consider a configuration of generalisation arcs and evidence to express abductive inference.

Definition 8 (Abductive inference)**.** *Let* $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ *be an IG, and let* $\mathbf{E} \subseteq \mathbf{P}$ *be an evidence set. Let* $p_1, \ldots, p_n, q \in \mathbf{P}$*, with* $\{p_1, \ldots, p_n\} \cap \mathbf{E} = \emptyset$*. Then given* $\mathbf{E}$*,* propositions $p_1, \ldots, p_n$ are abductively inferred from q using a $g \colon \{p_1, \ldots, p_n\} \to q$ in $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$*, denoted* $q \twoheadrightarrow_g p_1; \ldots ; q \twoheadrightarrow_g p_n$*, iff:*

(1) $q \in \mathbf{E}$*, or;*
(2) q *is deductively inferred from propositions* $r_1, \ldots, r_m \in \mathbf{P}$ *using a generalisation* $g' \colon \{r_1, \ldots, r_m\} \to q$ *in* $\mathbf{G}$*,* $g \neq g'$ *(see Definition 7), where* $g' \notin \mathbf{G}^{\mathrm{c}}$ *if* $g \in \mathbf{G}^{\mathrm{c}}$ *and* $g' \notin \mathbf{G}^{\mathrm{a}}$ *if* $g \in \mathbf{G}^{\mathrm{a}}$*, or;*
(3) q *is abductively inferred from a proposition* $r \in \mathbf{P}$ *using a generalisation* $g' \colon \{q, r_1, \ldots, r_m\} \to r$ *in* $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$*,* $r_1, \ldots, r_m \in \mathbf{P}$.

In accordance with our assumptions stated in Section 2.2, abduction is defeasible and is modelled using only causal generalisations and abstractions. Following Console and Dupré [13] and Bex [4], we assume that abductive inference can be performed with both strict and defeasible abstractions, where such an inference is always defeasible as it concerns an inference from the more abstract consequent to a more specific antecedent (see Section 2.2). The condition $\{p_1, \ldots, p_n\} \cap \mathbf{E} = \emptyset$ ensures that abduction cannot be performed with a generalisation to infer its antecedents in case at least one of its antecedents is already observed. Furthermore, abductive inference can only be performed using a generalisation $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$ to infer its antecedents $\mathbf{Tails}(g)$ from its consequent $Head(g)$ in case $Head(g)$ has been affirmed in that either $Head(g)$ is observed (i.e. $Head(g) \in \mathbf{E}$), $Head(g)$ is deductively inferred, or $Head(g)$ is itself abductively inferred.

In correspondence with Pearl's constraint (see Section 2.4.1), we assume in condition 2 that propositions $p_1, \ldots, p_n \in \mathbf{P}$ cannot be abductively inferred from a proposition $q \in \mathbf{P}$ using a generalisation $g \in \mathbf{G}^{\mathrm{c}}$ if its consequent q is deductively inferred using a generalisation $g' \neq g$, $g' \in \mathbf{G}^{\mathrm{c}}$. In enforcing this constraint, we do not need to consider whether or not the antecedents of g or g' include enablers, as illustrated in Example 12 from Section 2.4.1. More specifically, in Definition 2 it is assumed that $\forall g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}}$, $\mathbf{Ant}(g) \neq \emptyset$; therefore, at least one proposition p_i is an actual antecedent of g and

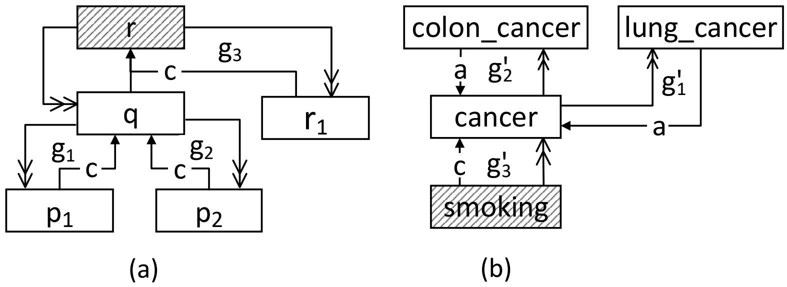


Fig. 10. Example of an IG illustrating abductive inference with causal generalisations (a); example of an IG illustrating abductive inference with abstractions (b).

at least one proposition r_j is an actual antecedent of g', which are then alternative explanations of q according to case 2a of Definition 3 which may not be inferred from each other by inferring q as an intermediary step. Similarly, we assume in condition 2 that $g' \notin \mathbf{G}^{\mathrm{a}}$ if $g \in \mathbf{G}^{\mathrm{a}}$ to account for our constraints on performing deduction and abduction in that order with two abstractions (see Section 2.4.2). In this case, propositions $p_1, \ldots, p_n \in \mathbf{Tails}(g)$ are alternative explanations of $r_1, \ldots, r_m \in \mathbf{Tails}(g')$ as indicated by g and g' according to case 3 of Definition 3.

Example 29. *In the IG of Figure 7, given* $\mathbf{E}$ *proposition* $155a$ *is abductively inferred from* 155 *using* $g_7 \in \mathbf{G}^{\mathrm{a}}_{\mathrm{d}}$*, as* 155 *is deductively inferred (Definition 8, condition* 2*). In turn, propositions* 156 *and* Π_3 *are iteratively abductively inferred using generalisations* $g_8 \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}}$ *and* $g_9 \in \mathbf{G}^{\mathrm{c}}$*, respectively. Note that although* g_8 *is a strict abstraction, the abductive inference from* $155a$ *to* 156 *is defeasible and not strict; specifically, that Sacco was conscious of having been involved in a robbery and shooting does not allow us to strictly infer that he was conscious of having been involved in the specific robbery and shooting that took place in South Braintree.*

In the IG of Figure 10a, q *and* r_1 *are abductively inferred from* r *using generalisation* $g_3 \colon \{q, r_1\} \rightarrow r$ *in* $\mathbf{G}^{\mathrm{c}}$ *by condition* 1 *of Definition 8, as* $r \in \mathbf{E}$*. Then by condition* 3 *of Definition 8,* p_1 *and* p_2 *are abductively inferred from* q *using* g_1 *and* g_2*, respectively.* □

The following example further illustrates abductive inference with abstractions.

Example 30. *In Figure 10b, Example 15 from Section 2.4.2 is modelled as an IG. As* smoking $\in \mathbf{E}$*,* cancer *is deductively inferred from* smoking *using* g'_3*. Propositions* lung_cancer *and* colon_cancer *are then abductively inferred from* cancer *using strict abstractions* g'_1 *and* g'_2*, respectively (Definition 8, condition* 2*). Hence, in this example, a cause (*smoking*) for an event (*cancer*) is known, after which this event is inferred and is in turn further specified at a lower level of abstraction (*lung_cancer *or* colon_cancer*). As noted in Section 2.4.2, this type of mixed inference using a causal generalisation and abstractions does not lead to undesirable results.* □

The following examples illustrate that Pearl's constraint for mixed deductive-abductive inference (see Section 2.4.1), as well as our proposed constraints on performing inference with abstractions (see Section 2.4.2), are adhered to.

Example 31. *In Figure 11a, Example 9 is modelled as an IG. As* smoke_machine $\in \mathbf{E}$*,* smoke *is deductively inferred from* smoke_machine *using* g_1*.* fire *cannot be inferred from* smoke*, as* $g_2 \in \mathbf{G}^{\mathrm{c}}$ *and* smoke *is deductively inferred using* $g_1 \in \mathbf{G}^{\mathrm{c}}$ *(Definition 8, condition* 2*).*

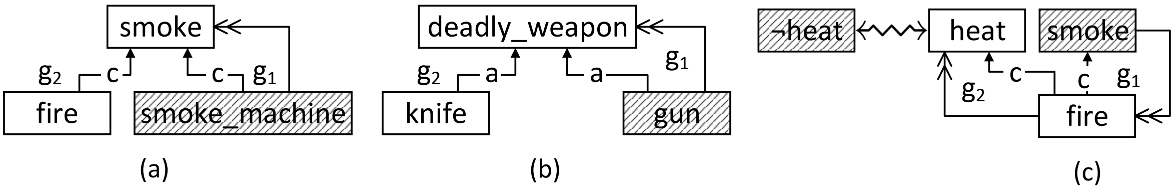


Fig. 11. An IG illustrating Pearl's constraint for mixed deductive-abductive inference (a); an IG illustrating our inference constraints for abstractions (b); an IG illustrating mixed abductive-deductive inference (c).

In Figure 11b, Example 13 is modelled as an IG. As gun $\in$ **E**, deadly_weapon *is deductively inferred from* gun *using* g_1. knife *cannot in turn be inferred from* deadly_weapon, *as* $g_2 \in \mathbf{G}^{\mathrm{a}}$ *and* deadly_weapon *is deductively inferred using* $g_1 \in \mathbf{G}^{\mathrm{a}}$ *(Definition 8, condition* 2*).* □

The following example describes the inferences that can be made based on the IG of Figure 5 corresponding to the mind map example of Section 3.2.

Example 32. *Consider the IG of Figure 12. Given* $\mathbf{E} = \{$tes$_1$, tes$_2$, tes$_3$, autopsy$\}$, head_wound *is deductively inferred from* autopsy *using* g_5*. Then,* hit_angular *and* fell_on_table *are abductively inferred from* head_wound *using* g_6 *and* g_7*, respectively (Definition 8, condition* 2*).* head_wound *is also deductively inferred from* fell_on_table *using* g_7*, as* fell_on_table *is deductively inferred from* tes$_3$ *using* g_8*; the inference type of* g_7 *is, therefore, ambiguous (see Section 2.5).* hammer *and* stone *are abductively inferred from* hit_angular *using* g_3 *and* g_4*, respectively (Definition 8, condition* 3*).* hit_angular *is also deductively inferred from* hammer *and* stone *using* g_3 *and* g_4*, respectively, as* hammer *is deductively inferred from* tes$_1$ *using* g_1 *and* stone *is deductively inferred from* tes$_2$ *using* g_2*. Then,* head_wound *is deductively inferred from* hit_angular *using* g_6*.* □

4.2.3. Mixed abductive-deductive inference

As apparent from Definitions 7 and 8, mixed abductive-deductive inference can be performed within our IG-formalism.

Example 33. *In Figure 11c, Example 7 from Section 2.4 is modelled as an IG. From* smoke*,* fire *is abductively inferred using* g_1*, as* smoke $\in$ **E**. *Then* heat *is deductively inferred (or predicted) from* fire *using* g_2 *(Definition 7, condition* 3*).* □

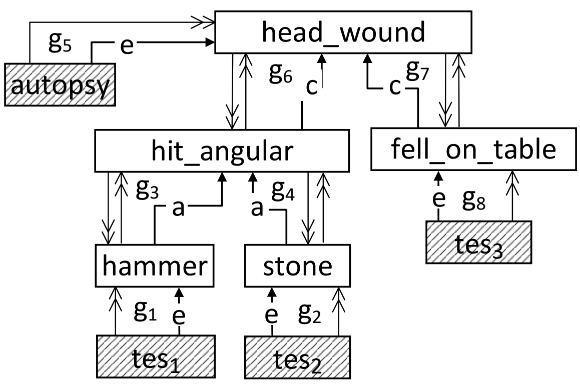


Fig. 12. The IG of Figure 5, where evidence set **E** (shaded) and resulting inference steps (↠) are also indicated.

5. An argumentation formalism based on information graphs

Based on our IG-formalism from Section 4, we now define an argumentation formalism that allows for both deductive and abductive argumentation. Note that the IG-formalism is not an argumentation formalism, and that no semantics for IGs were defined in Section 4. Instead, we defined how inference can be performed with IGs and we defined different notions of conflicts. In the current section, we define an argumentation formalism based on IGs which allows us to assign a semantics to argumentation frameworks constructed on the basis of IGs. More specifically, our approach generates an abstract argumentation framework as in Dung [14], that is, a set of arguments with a binary attack relation, which thus allows arguments based on IGs to be formally evaluated according to Dung's semantics. We can then study properties of generated AFs; in particular, we prove that Caminada and Amgoud's [10] postulates are satisfied by instantiations of our formalism, which warrants the sound definition of instantiations of our argumentation system and implies that anomalous results such as issues regarding inconsistency and non-closure as identified by [10] are avoided. Our argumentation formalism extends a preliminary version proposed in [38] that was based on a more restricted version of our IG-formalism [39] in which only causal and evidential generalisations without enablers were considered. Moreover, satisfaction of rationality postulates was not proven in that paper.

In Section 5.1, we define arguments on the basis of a provided IG and an evidence set $\mathbf{E}$, which capture sequences of deductive and abductive inference applications as defined in Definitions 7 and 8 starting with elements from $\mathbf{E}$. We then formally prove that arguments constructed on the basis of IGs conform to our inference constraints (Section 2.4). In Section 5.2, we define several types of attacks between arguments based on IGs, which are based on the different types of conflicts defined for our IG-formalism. In Section 5.3 we instantiate Dung's abstract approach with arguments and attacks based on IGs and provide the definitions of Dung's argumentation semantics. In Section 5.4, we then prove that rationality postulates [10] are satisfied by instantiations of our formalism.

5.1. Arguments

In this section, we define how arguments on the basis of an IG and an evidence set $\mathbf{E}$ are constructed. Here, we take inspiration from the definition of an argument as defined for the ASPIC$^+$ framework [21]. By remaining close to the ASPIC$^+$ framework, this allows us to straightforwardly show that rationality postulates are satisfied for our argumentation formalism based on IGs (see Section 5.4). In what follows, for a given argument, the operator PREM returns all propositions in $\mathbf{E}$ used to construct the argument, CONC returns its conclusion, SUB returns all its sub-arguments (including itself), IMMSUB returns its immediate sub-arguments, GEN returns all the generalisations used in constructing the argument, TOPGEN returns the last generalisation used in constructing the argument, DEFINF and STINF return all the defeasible and strict inferences used in constructing the argument, respectively, and TOPINF returns the last inference used in constructing the argument. Definition 9 is explained and illustrated in Examples 34 and 35.

Definition 9 (Argument)**.** *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG, and let $\mathbf{E} \subseteq \mathbf{P}$ be an evidence set. An* argument A on the basis of $G_{\mathcal{I}}$ and $\mathbf{E}$ *is any structure obtainable by applying one or more of the following steps finitely many times, where steps* 2 *(i.e. step* 2a *or* 2b*) and* 3 *or vice versa are not subsequently applied using the same generalisation arc $g \in \mathbf{G}$:*

1. *p if $p \in \mathbf{E}$, where:* $\textsc{Prem}(A) = \{p\}$*;* $\textsc{Conc}(A) = p$*;* $\textsc{Sub}(A) = \{A\}$*;* $\textsc{ImmSub}(A) = \emptyset$*;* $\textsc{Gen}(A) = \emptyset$*;* $\textsc{TopGen}(A) =$ *undefined;* $\textsc{DefInf}(A) = \emptyset$*;* $\textsc{StInf}(A) = \emptyset$*;* $\textsc{TopInf}(A) =$ *undefined.*

2a. *$A_1, \ldots, A_n \twoheadrightarrow_g p$ if $A_1, \ldots, A_n$ are arguments such that p is defeasibly deductively inferred from $\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)$ using a generalisation $g \colon \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)\} \rightarrow p$ according to Definition 7, where it holds that $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$ and if g is of the form $g \colon c \rightarrow e$ in $\mathbf{G}^{\mathrm{c}}$ and its evidential counterpart $g' \colon e \rightarrow c$ is included in $\mathbf{G}^{\mathrm{e}}$, then $g' \notin \textsc{Gen}(A_1) \cup \ldots \cup \textsc{Gen}(A_n)$. For A, it holds that:*
$\textsc{Prem}(A) = \textsc{Prem}(A_1) \cup \ldots \cup \textsc{Prem}(A_n)$*;* $\textsc{Conc}(A) = p$*;*
$\textsc{Sub}(A) = \textsc{Sub}(A_1) \cup \ldots \cup \textsc{Sub}(A_n) \cup \{A\}$*;* $\textsc{ImmSub}(A) = \{A_1, \ldots, A_n\}$*;*
$\textsc{Gen}(A) = \textsc{Gen}(A_1) \cup \ldots \cup \textsc{Gen}(A_n) \cup \{g\}$*;* $\textsc{TopGen}(A) = g$*;*
$\textsc{DefInf}(A) = \textsc{DefInf}(A_1) \cup \ldots \textsc{DefInf}(A_n) \cup \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n) \twoheadrightarrow_g p\}$*;*
$\textsc{StInf}(A) = \textsc{StInf}(A_1) \cup \ldots \textsc{StInf}(A_n)$*;*
$\textsc{TopInf}(A) = \textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n) \twoheadrightarrow_g p$*.*

2b. *$A_1, \ldots, A_n \rightharpoonup_g p$ if $A_1, \ldots, A_n$ are arguments such that p is strictly deductively inferred from $\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)$ using a generalisation $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$, $g \colon \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)\} \rightarrow p$ according to Definition 7, where $\textsc{Prem}(A)$, $\textsc{Conc}(A)$, $\textsc{Sub}(A)$, $\textsc{ImmSub}(A)$, $\textsc{Gen}(A)$ and $\textsc{TopGen}(A)$ are defined as in step 2a, and where:*
$\textsc{DefInf}(A) = \textsc{DefInf}(A_1) \cup \ldots \textsc{DefInf}(A_n)$*;*
$\textsc{StInf}(A) = \textsc{StInf}(A_1) \cup \ldots \textsc{StInf}(A_n) \cup \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n) \rightharpoonup_g p\}$*;*
$\textsc{TopInf}(A) = \textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n) \rightharpoonup_g p$*.*

3. *$A' \twoheadrightarrow_g p$ if A' is an argument such that p is abductively inferred from $\textsc{Conc}(A')$ using a generalisation $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$, $g \colon \{p, p_1, \ldots, p_n\} \rightarrow \textsc{Conc}(A')$ for some propositions $p_1, \ldots, p_n \in \mathbf{P}$ according to Definition 8, where:*
$\textsc{Prem}(A) = \textsc{Prem}(A')$*;* $\textsc{Conc}(A) = p$*;* $\textsc{Sub}(A) = \textsc{Sub}(A') \cup \{A\}$*;* $\textsc{ImmSub}(A) = \{A'\}$*;* $\textsc{Gen}(A) = \textsc{Gen}(A') \cup \{g\}$*;* $\textsc{TopGen}(A) = g$*;* $\textsc{DefInf}(A) = \textsc{DefInf}(A') \cup \{\textsc{Conc}(A') \twoheadrightarrow_g p\}$*;* $\textsc{StInf}(A) = \textsc{StInf}(A')$*;* $\textsc{TopInf}(A) = \textsc{Conc}(A') \twoheadrightarrow_g p$*.*

Note that we overload symbols $\twoheadrightarrow$ and $\rightharpoonup$ to denote an argument while it also denotes a defeasible or strict inference. The set of all arguments on the basis of $G_{\mathcal{I}}$ and $\mathbf{E}$ is denoted by $\mathcal{A}$.

An argument $A \in \mathcal{A}$ is called *strict* if $\textsc{DefInf}(A) = \emptyset$; otherwise, A is called *defeasible*. An argument $A \in \mathcal{A}$ is called a *premise argument* if only step 1 of Definition 9 is applied, *deductive* if only steps 1, 2a and 2b are applied, *abductive* if only steps 1 and 3 are applied, and *mixed* otherwise. The restriction that steps 2 (i.e. step 2a or 2*b*) and 3 or vice versa are not subsequently applied using the same generalisation arc $g \in \mathbf{G}$ ensures that cycles in which two propositions are iteratively deductively and abductively inferred from each other using the same g are avoided in argument construction. Similarly, in case causal generalisation $g \colon c \rightarrow e$ has an evidential counterpart $g' \colon e \rightarrow c$ (see Sections 2.3 and 4.1), then the restriction in step 2a that $g' \notin \textsc{Gen}(A_1) \cup \ldots \cup \textsc{Gen}(A_n)$ ensures that cycles in which c and e are iteratively deductively inferred from each other using g' and g are avoided. Note that cycles in which c and e are iteratively deductively inferred from each other using g and g' in that order are already avoided due to the enforcement of Pearl's constraint in condition 2 of Definition 7.

Example 34. *Consider Figure 13, in which arguments constructed on the basis of the IG of Figure 7 are indicated. According to step* 1 *of Definition 9, $A_1 \colon 150$ and $A_2 \colon 151$ are premise arguments. Based on A_1 and A_2, defeasible deductive argument $A_3 \colon A_1, A_2 \twoheadrightarrow_{g_1} 149$ is constructed by step 2a of Definition 9, as 149 is defeasibly deductively inferred from 150 and 151 using $g_1 \in \mathbf{G}^{\mathrm{e}}$. Arguments $A_4 \colon A_3$*

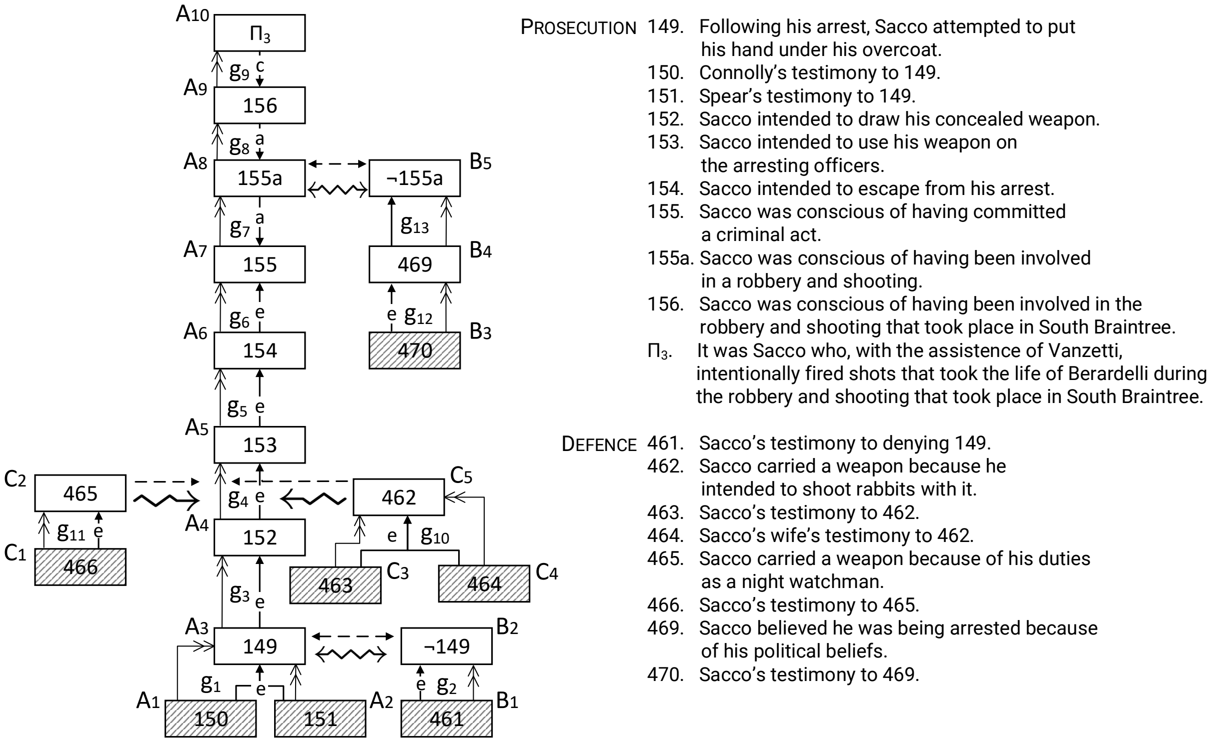


Fig. 13. The IG of Figure 7, where arguments and direct attacks (⇢) on the basis of the IG and $\mathbf{E}$ are also indicated.

$\twoheadrightarrow_{g_3} 152$; $A_5\colon A_4 \twoheadrightarrow_{g_4} 153$; $A_6\colon A_5 \twoheadrightarrow_{g_5} 154$ *and* $A_7\colon A_6 \twoheadrightarrow_{g_6} 155$ *similarly are defeasible deductive arguments. Argument* $A_8\colon A_7 \twoheadrightarrow_{g_7} 155a$ *is a defeasible mixed argument by step* 3 *of Definition 9, as* $155a$ *is abductively inferred from* 155 *using* g_7*. Similarly, arguments* $A_9\colon A_8 \twoheadrightarrow_{g_8} 156$ *and* $A_{10}\colon A_9 \twoheadrightarrow_{g_9} \Pi_3$ *are defeasible mixed arguments. To illustrate the operators used in Definition 9, for* A_8*, we have that* $\textsc{Prem}(A_8) = \{150, 151\}$; $\textsc{Conc}(A_8) = 155a$; $\textsc{Sub}(A_8) = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8\}$; $\textsc{ImmSub}(A_8) = \{A_7\}$; $\textsc{Gen}(A_8) = \{g_1, g_3, g_4, g_5, g_6, g_7\}$; $\textsc{TopGen}(A_8) = g_7$; $\textsc{DefInf}(A_8) = \{150, 151 \twoheadrightarrow_{g_1} 149; 149 \twoheadrightarrow_{g_3} 152; 152 \twoheadrightarrow_{g_4} 153; 153 \twoheadrightarrow_{g_5} 154; 154 \twoheadrightarrow_{g_6} 155; 155 \twoheadrightarrow_{g_7} 155a\}$; $\textsc{StInf}(A_8) = \emptyset$; $\textsc{TopInf}(A_8) = 155 \twoheadrightarrow_{g_7} 155a$. □

Step 3 of Definition 9 is now illustrated in more detail.

Example 35. *On the basis of the IG of Figure 10a and* $\mathbf{E} = \{r\}$, $A'_1\colon r$ *is a premise argument. From* A'_1*, arguments* $A'_2\colon A'_1 \twoheadrightarrow_{g_3} r_1$ *and* $A'_3\colon A'_1 \twoheadrightarrow_{g_3} q$ *are constructed by step* 3 *of Definition 9, as* q *and* r_1 *are abductively inferred from* $\textsc{Conc}(A'_1)$ *using causal generalisation* $g_3\colon \{q, r_1\} \to r$*. Then again by step* 3*,* $A'_4\colon A'_3 \twoheadrightarrow_{g_1} p_1$ *and* $A'_5\colon A'_3 \twoheadrightarrow_{g_2} p_2$ *are constructed using* g_1 *and* g_2*, respectively.* □

5.1.1. Properties of arguments based on IGs

We now prove a number of formal properties of arguments based on IGs. Lemma 1 states that the conclusions of deductive, abductive, and mixed arguments constructed in our argumentation formalism based on IGs are not observed.

Lemma 1. *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Let $A \in \mathcal{A}$ be a deductive, abductive, or mixed argument. Then* $\text{CONC}(A) \notin \mathbf{E}$.

Proof. As A is not a premise argument, step 2a, step 2b or step 3 of Definition 9 is applied last in constructing A. In case step 2a or 2b of Definition 9 is applied last, then $\exists g \in \mathbf{G}$ such that $Head(g) = \text{CONC}(A)$ is deductively inferred using $\text{TOPGEN}(A) = g$ according to Definition 7. Hence, per the restrictions of Definition 7, $Head(g) = \text{CONC}(A) \notin \mathbf{E}$. In case step 3 of Definition 9 is applied last, then $\exists g \in \mathbf{G}$ such that $\text{CONC}(A) \in \mathbf{Tails}(g)$ is abductively inferred using $\text{TOPGEN}(A) = g$ according to Definition 8. Hence, $\text{CONC}(A) \notin \mathbf{E}$ per the restriction of Definition 8 that $\mathbf{Tails}(g) \cap \mathbf{E} = \emptyset$. ■

In performing inference care should be taken that no cause for an effect is inferred in case an alternative cause for this effect was already previously inferred (see Section 2.4.1). Similarly, care should be taken that no version of an event at a lower level of abstraction is inferred if an alternative version of this event at a lower level of abstraction was already previously inferred (see Section 2.4.2). In the context of IGs, for $g \in \mathbf{G}^{\mathrm{c}}$, propositions in $\mathbf{Ant}(g)$ express causes for the common effect expressed by $Head(g)$, for $g \in \mathbf{G}^{\mathrm{e}}$, $Head(g)$ expresses the usual cause for propositions in $\mathbf{Ant}(g)$, and for $g \in \mathbf{G}^{\mathrm{a}}$, propositions in $\mathbf{Tails}(g)$ are at a lower level of abstraction than $Head(g)$. Hence, in defining how inferences can be read from IGs, restrictions are put in Definitions 7 and 8 such that our inference constraints (see Section 2.4) are adhered to. We now formally prove that these inference constraints are never violated in constructing sequences of arguments on the basis of IGs.

First, we formally define the inference constraints of Section 2.4 in the context of arguments constructed on the basis of IGs.

Definition 10 (Inference constraint)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Let $p_1, p_2 \in \mathbf{P}$ be alternative explanations of $q \in \mathbf{P}$ as indicated by generalisations $g_1, g_2 \in \mathbf{G}$ (Definition 3). If arguments A and B exist in $\mathcal{A}$ with $\text{CONC}(B) = q$, $A \in \text{IMMSUB}(B)$, and $\text{CONC}(A) = p_1$, then there does not exist an argument $C \in \mathcal{A}$ with $B \in \text{IMMSUB}(C)$, $\text{CONC}(C) = p_2$.*

We now formally prove that this inference constraint is indeed adhered to.

Proposition 1 (Adherence to inference constraint)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Then $\mathcal{A}$ adheres to the inference constraint of Definition 10.*

Proof. Assume that $p_1, p_2 \in \mathbf{P}$ are alternative explanations of $q \in \mathbf{P}$ as indicated by generalisations g_1 and g_2 in $\mathbf{G}$, and assume that arguments $A, B \in \mathcal{A}$ exist with $\text{CONC}(B) = q$, $A \in \text{IMMSUB}(B)$, $\text{CONC}(A) = p_1$. Then we need to prove that no argument C exists in $\mathcal{A}$ with $B \in \text{IMMSUB}(C)$ and $\text{CONC}(C) = p_2$. In constructing argument B, either step 2a, step 2b or step 3 of Definition 9 is applied last, where generalisation g_1 is used to infer $\text{CONC}(B) = q$. Here, g_1 cannot be of the form $g_1 \in \mathbf{G}^{\mathrm{e}}$, $q \in \mathbf{Ant}(g_1)$, $Head(g_1) = p_1$ (Definition 3, case 1) as in this case antecedent q of g_1 is inferred from consequent p_1 of g_1, which would be an instance of abductive inference while per the restrictions of Definition 8 abductive inference can only be performed using generalisations in $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$. More specifically, argument B cannot be constructed by applying step 2a, 2b and 3 of Definition 9 last if g_1 is of that form. Thus, we only need to consider cases 2 and 3 of Definition 3, where a generalisation $g_1 \in \mathbf{G}^{\mathrm{c}}$, $Head(g_1) = q$, $p_1 \in \mathbf{Ant}(g_1)$ respectively a generalisation $g_1 \in \mathbf{G}^{\mathrm{a}}$, $Head(g_1) = q$, $p_1 \in \mathbf{Tails}(g_1)$ is used to construct B, namely by applying step 2a or 2b of Definition 9 last to deductively infer $\text{CONC}(B) = q$. We now show that for the given options for g_1, no argument C with $B \in \text{IMMSUB}(C)$, $\text{CONC}(C) = p_2$ can be constructed using g_2.

- First, consider case 2a of Definition 3 in which $g_2 \neq g_1$, $g_2 \in \mathbf{G}^{\mathrm{c}}$, $Head(g_2) = q$, $p_2 \in \mathbf{Ant}(g_2)$. Then no argument C with $B \in \textsc{ImmSub}(C)$, $\textsc{Conc}(C) = p_2$ can be constructed using g_2, as in this case abduction would be performed with g_2 to infer p_2 from q while per the restrictions in condition 2 of Definition 8 abduction cannot be performed with g_2 as $Head(g_2)$ was previously deductively inferred using $g_1 \in \mathbf{G}^{\mathrm{c}}$. In particular, step 3 of Definition 9 cannot be applied in constructing C using g_2. Furthermore, neither step 2a nor step 2b of Definition 9 can be applied in constructing C using g_2, as these steps specify deductive and not abductive inferences.
- Next, consider case 2b of Definition 3 in which $g_2 \in \mathbf{G}^{\mathrm{e}}$, $Head(g_2) = p_2$, $q \in \mathbf{Ant}(g_2)$. Then no argument C with $B \in \textsc{ImmSub}(C)$, $\textsc{Conc}(C) = p_2$ can be constructed using g_2, as in this case deductive inference would be performed with g_2 to infer p_2 while per the restrictions in condition 2 of Definition 7 deductive inference cannot be performed with g_2 as $q \in \mathbf{Ant}(g_2)$ was previously deductively inferred using $g_1 \in \mathbf{G}^{\mathrm{c}}$. In particular, step 2a of Definition 9 cannot be applied in constructing C using g_2. Furthermore, step 2b of cannot be applied in constructing C using g_2, as this step can only be applied using strict generalisations and $g_2 \notin \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$, and step 3 cannot be applied in constructing C using g_2, as this step specifies an abductive and not a deductive inference.
- Finally, consider case 3 of Definition 3 in which $g_2 \neq g_1$, $g_2 \in \mathbf{G}^{\mathrm{a}}$, $Head(g_2) = q$, $p_2 \in \mathbf{Tails}(g_2)$. Then no argument C with $B \in \textsc{ImmSub}(C)$, $\textsc{Conc}(C) = p_2$ can be constructed using g_2, as in this case abduction would be performed with g_2 to infer p_2 from q while per the restrictions in condition 2 of Definition 8 abduction cannot be performed with g_2 as $Head(g_2)$ was previously deductively inferred using $g_1 \in \mathbf{G}^{\mathrm{a}}$. In particular, step 3 of Definition 9 cannot be applied in constructing C using g_2. Furthermore, neither step 2a nor step 2b of Definition 9 can be applied in constructing C using g_2, as these steps specify deductive and not abductive inferences. ■

5.2. Attack

In this section, several types of attacks between arguments on the basis of IGs are defined. Among the types of attacks that are typically distinguished in structured argumentation (for instance in ASPIC$^+$ [21]) are rebuttal, undermining, and undercutting attack. Of these types of attacks, we only consider rebuttal and undercutting attack and not undermining attacks (i.e. attack on an argument's premises [21]), as in IGs we assume that all premises are certain and cannot be attacked (cf. ASPIC$^+$'s axiom premises). We also distinguish a fourth type of attack, namely alternative attack, a concept based on the notion of competing alternative explanations (see Section 2.2) that is inspired by [3, 5]. In our argumentation formalism, attacks directly follow from the constructed arguments and the specified exception arcs in an IG. Hence, attacks between arguments do not need to be separately specified by the user.

First, we define the general notion of attack, after which the different types of attacks are defined.

Definition 11 (Attack)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}}$ and evidence set $\mathbf{E}$. Let $A, B \in \mathcal{A}$. Then A* attacks *B iff A rebuts B, A undercuts B, or A alternative attacks B, as defined in Definitions 12, 13 and 14, respectively.*

5.2.1. Rebuttal attack

First, rebuttal attack is defined. Informally, a rebuttal is an attack on the conclusion of an argument for which it holds that the last inference used in constructing the argument is defeasible.

Definition 12 (Rebuttal attack)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Let A, B and B' be arguments in $\mathcal{A}$ with $B' \in \text{SUB}(B)$. Then A rebuts B (on B') iff there exists a negation arc n: $\text{CONC}(A) \leftrightsquigarrow \text{CONC}(B')$ in $\mathbf{N}$ and B' is of the form $B''_1, \ldots, B''_n \twoheadrightarrow_g p$ for some $B''_1, \ldots, B''_n \in \mathcal{A}$, $p \in \mathbf{P}$.*

Note that, as it is assumed that B' is of the form $B''_1, \ldots, B''_n \twoheadrightarrow_g p$ (i.e. $\text{TOPINF}(B')$ is defeasible), it holds that B' is a deductive, abductive, or mixed argument; hence, by Lemma 1, $\text{CONC}(B') \notin \mathbf{E}$. Furthermore, while a negation arc expresses a symmetric conflict, our definition of rebuttal attack allows for both symmetric or asymmetric rebuttal, as illustrated by the following example.

Example 36. *Consider the IG of Figure 13. Let A_1, A_2, A_3 be the arguments introduced in Example 34. Let B_1: 461 and let B_2: $B_1 \twoheadrightarrow_{g_2} \neg 149$. Then B_2 rebuts A_3 (on A_3) and A_3 rebuts B_2 (on B_2), as $\text{CONC}(A_3) = 149$, $\text{CONC}(B_2) = \neg 149$ (and hence n: $149 \leftrightsquigarrow \neg 149$ in $\mathbf{N}$), where $\text{TOPINF}(A_3) = 150, 151 \twoheadrightarrow_{g_1} 149$ and $\text{TOPINF}(B_2) = 461 \twoheadrightarrow_{g_2} \neg 149$ are defeasible. This symmetric rebuttal is indicated in Figure 13 by means of a bidirectional dashed arc between these propositions. Similarly, let A_8 be as introduced in Example 34, and let B_3: 470; B_4: $B_3 \twoheadrightarrow_{g_{12}} 469$; B_5: $B_4 \twoheadrightarrow_{g_{13}} \neg 155a$. Then A_8 rebuts B_5 (on B_5) and B_5 rebuts A_8 (on A_8).*

Consider again Example 33, in which heat *is predicted from* fire*. Assume that contrary to this prediction we observe that there is no heat ($\neg$heat $\in \mathbf{E}$). Let A''_1: smoke; A''_2: $A''_1 \twoheadrightarrow_{g_1}$ fire; A''_3: $A''_2 \twoheadrightarrow_{g_2}$ heat; B''_1: $\neg$heat. Then B''_1 rebuts A''_3 (on A''_3), but A''_3 does not rebut B''_1 as B''_1 is not of the form $B'''_1, \ldots, B'''_n \twoheadrightarrow_g p$ for some $B'''_1, \ldots, B'''_n \in \mathcal{A}$, $p \in \mathbf{P}$ (i.e. B''_1 is a premise argument).* □

5.2.2. Undercutting attack

Next, undercutting attack is considered. Informally, an undercutter attacks a defeasible inference by providing exceptional circumstances under which the inference may not be applicable. In our argumentation formalism based on IGs, undercutting attacks between arguments follow from the specified exception arcs in $G_{\mathcal{I}}$. Specifically, as an exception arc directed from $p \in \mathbf{P}$ to $g \in \mathbf{G}^{\text{c}} \cup \mathbf{G}^{\text{e}} \cup \mathbf{G}^{\text{a}}_{\text{d}} \cup \mathbf{G}^{\text{o}}_{\text{d}}$ specifies an exception to defeasible generalisation g, an argument $A \in \mathcal{A}$ with $\text{CONC}(A) = p$ undercuts an argument $B \in \mathcal{A}$ with $g \in \text{GEN}(B)$.

Definition 13 (Undercutting attack)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Let $A, B, B' \in \mathcal{A}$ with $B' \in \text{SUB}(B)$. Then A* undercuts *B (on B') iff there exists an exception arc $x \in \mathbf{X}$ such that x: $\text{CONC}(A) \rightsquigarrow g$ and $\text{TOPGEN}(B') = g \in \mathbf{G}^{\text{c}} \cup \mathbf{G}^{\text{e}} \cup \mathbf{G}^{\text{a}}_{\text{d}} \cup \mathbf{G}^{\text{o}}_{\text{d}}$.*

Undercutting attack is illustrated by the following example.

Example 37. *Consider the IG of Figure 13. Let A_1, A_2, A_3, A_4, A_5 be the arguments introduced in Example 34. Let C_1: 466; C_2: $C_1 \twoheadrightarrow_{g_{11}} 465$. Then C_2 undercuts A_5 (on A_5), as x: $465 \rightsquigarrow g_4$ in $\mathbf{X}$ and $\text{TOPGEN}(A_5) = g_4 \in \mathbf{G}^{\text{e}}$. This direct attack is indicated in Figure 13 by means of a dashed arc directed from 465 to defeasible inference $152 \twoheadrightarrow_{g_4} 153$. As undercutting attack is defined on subarguments, C_2 also attacks A_i for $i \geqslant 6$. Similarly, let C_3: 463; C_4: 464; C_5: $C_3, C_4 \twoheadrightarrow_{g_{10}} 462$. Then C_5 undercuts A_5 (on A_5), as x: $462 \rightsquigarrow g_4$ in $\mathbf{X}$ and $\text{TOPGEN}(A_5) = g_4$. Argument C_5 then also attacks A_i for $i \geqslant 6$.* □

5.2.3. Alternative attack

Lastly, alternative attack is defined. Arguments are involved in alternative attack iff their abductively inferred conclusions are competing alternative explanations (see Section 2.2).

Definition 14 (Alternative attack)**.** *Let $\mathcal{A}$ be a set of arguments on the basis of IG $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ and evidence set $\mathbf{E}$. Let $p_1, p_2 \in \mathbf{P}$ be alternative explanations of $q \in \mathbf{P}$ as indicated by generalisations g and g' in $\mathbf{G}$, where either $g, g' \in \mathbf{G}^{\mathrm{c}}$ (Definition 3, case 2a) or $g, g' \in \mathbf{G}^{\mathrm{a}}$ (Definition 3, case 3). Let $A, B, B' \in \mathcal{A}$ with $B' \in \textsc{Sub}(B)$. Then A alternative attacks B (on B') iff there exists an argument $C \in \textsc{ImmSub}(A) \cap \textsc{ImmSub}(B')$ such that $\textsc{Conc}(A) = p_1$ and $\textsc{Conc}(B') = p_2$ are abductively inferred from $\textsc{Conc}(C) = q$ using generalisations g and g', respectively.*

Note that A only alternative attacks B on B' iff $\textsc{TopInf}(B')$ is an abductive inference and hence iff the last used inference in constructing B' is defeasible. Furthermore, unlike direct rebuttal attack, which can either be symmetric or asymmetric, direct alternative attack is always symmetric in that A alternative attacks B on B iff B alternative attacks A on A.

Under the conditions set out in Definition 14, arguments $A_i\colon C \twoheadrightarrow_g p_i$ for $p_i \in \mathbf{Ant}(g)$ constructed from C via abductive inference using g are involved in alternative attack with $A'_j\colon C \twoheadrightarrow_{g'} p'_j$ for $p'_j \in \mathbf{Ant}(g')$ constructed from C via abductive inference using g'. We do not consider arguments $A_i\colon C \twoheadrightarrow_g p_i$ for $p_i \in \mathbf{Enabler}(g)$ to be in competition with arguments $A'_j\colon C \twoheadrightarrow_{g'} p'_j$ for $p'_j \in \mathbf{Enabler}(g')$, as enablers of causal generalisations do not express alternative causes for the consequent. Arguments A_i (as well as A'_j) are not involved in alternative attack *among themselves*, in accordance with our assumption that the antecedents of a causal generalisation or abstraction are not in competition. Finally, in case $g \in \mathbf{G}^{\mathrm{c}}$ and $g' \in \mathbf{G}^{\mathrm{a}}$, then arguments A_i are not involved in alternative attack with A'_j, as the actual antecedents of g express causes for the effect expressed by the consequent but the tails of g' are not alternative explanations of the consequent; instead, propositions in $\mathbf{Tails}(g')$ are at a lower level of abstraction than $Head(g')$.

Example 38. *Consider the IG of Figure 12. Given $\mathbf{E}$, arguments $D_1\colon$* autopsy*; $D_2\colon D_1 \twoheadrightarrow_{g_5}$* head_wound*; $D_3\colon D_2 \twoheadrightarrow_{g_6}$* hit_angular*; $D_4\colon D_2 \twoheadrightarrow_{g_7}$* fell_on_table*; $D_5\colon D_3 \twoheadrightarrow_{g_3}$* hammer*; and $D_6\colon D_3 \twoheadrightarrow_{g_4}$* stone *are constructed. Here,* hit_angular *and* fell_on_table *are abductively inferred from* head_wound *using g_6 and g_7, respectively, and* hammer *and* stone *are abductively inferred from* hit_angular *using g_3 and g_4, respectively. Then D_3 alternative attacks D_4 (on D_4) and D_4 alternative attacks D_3 (on D_3), as $\textsc{Conc}(D_3) =$* hit_angular *and $\textsc{Conc}(D_4) =$* fell_on_table *are alternative explanations of $\textsc{Conc}(D_2) =$* head_wound *as indicated by g_6 and g_7 in $\mathbf{G}^{\mathrm{c}}$ (Definition 3, case 2a). As $D_3 \in \textsc{Sub}(D_5)$ and $D_3 \in \textsc{Sub}(D_6)$, D_4 also alternative attacks D_5 and D_6 (on D_3). Finally, D_5 alternative attacks D_6 (on D_6) and D_6 alternative attacks D_5 (on D_5), as $\textsc{Conc}(D_5) =$* hammer *and $\textsc{Conc}(D_6) =$* stone *are alternative explanations of $\textsc{Conc}(D_3) =$* hit_angular *as indicated by g_3 and g_4 in $\mathbf{G}^{\mathrm{a}}$ (Definition 3, case 3).*

Consider Example 12 from Section 2.4.1. Assume that in addition to generalisations g_1 and g_2, evidential generalisation $g_3\colon$ see_fire → fire *is provided. Given $\mathbf{E} = \{$*see_fire*$\}$, arguments $E_1\colon$* see_fire*; $E_2\colon E_1 \twoheadrightarrow_{g_3}$* fire*; $E_3\colon E_2 \twoheadrightarrow_{g_1}$* torch*; $E_4\colon E_2 \twoheadrightarrow_{g_2}$* match*; and $E_5\colon E_2 \twoheadrightarrow_{g_2}$* oxygen *are constructed. Then E_3 and E_4 are involved in alternative attack, as $\textsc{Conc}(E_3) =$* torch *and $\textsc{Conc}(E_4) =$* match *are alternative explanations of $\textsc{Conc}(E_2) =$* fire *as indicated by g_1 and g_2 in $\mathbf{G}^{\mathrm{c}}$ (Definition 3, case 2a), where* torch *and* match *are abductively inferred from* fire *using g_1 and g_2, respectively. E_3 is not involved in alternative attack with E_5, as $\textsc{Conc}(E_5) =$* oxygen $\in \mathbf{Enabler}(g_2)$.

Consider Figure 10a. Let A'_1, A'_2, A'_3 be as defined in Example 35. Then A'_2 and A'_3 are not *involved in alternative attack, as $r_1 = \textsc{Conc}(A'_2)$ and $q = \textsc{Conc}(A'_3)$ are abductively inferred from $r = \textsc{Conc}(A'_1)$ using the* same *generalisation g_3; specifically, in case 2a of Definition 3 it is assumed that $g \neq g'$, and hence r_1 and q are not alternative explanations of r by that definition.* □

Finally, note that a causal generalisation $g_1\colon c_1 \to e$ may be replaced by an evidential generalisation $g_1'\colon e \to c_1$ if c_1 is the usual cause of e, in which case abductive inference with g_1 can be encoded as deductive inference with g_1' (see Section 2.3). Considering the case in which only g_1 and not g_1' is included in IG $G_\mathcal{I}$ and additional causal generalisation $g_2\colon c_2 \to e$ is provided, then arguments $A_1\colon e$, $A_2\colon A_1 \twoheadrightarrow_{g_1} c_1$, $A_3\colon A_1 \twoheadrightarrow_{g_2} c_2$ are constructed upon observing e, where A_2 and A_3 are involved in alternative attack according to Definition 14. However, in case only g_1' and g_2 are included in $G_\mathcal{I}$ and not g_1, then arguments $A_1\colon e$, $A_2'\colon A_1 \twoheadrightarrow_{g_1'} c_1$, $A_3\colon A_1 \twoheadrightarrow_{g_2} c_2$ are constructed, where A_2' and A_3 are not involved in alternative attack as $g_1' \in \mathbf{G}^{\mathrm{e}}$. Hence, if the knowledge engineer considers c_1 and c_2 to be competing alternative explanations of e, then the involved generalisations should be modelled as causal generalisations in order to achieve alternative attack among constructed arguments. Alternatively, A_3 can be interpreted as an undercutter of A_2' as it provides an exception to the performed inference (see also [5, p. 15]). We reiterate that it is the responsibility of the knowledge engineer in consultation with the domain expert to decide which knowledge (including conflicts) to represent in an IG and to ensure this knowledge is modelled correctly (see also Section 4.1).

5.3. Argument evaluation

In this section, we provide Dung's definitions for argumentation semantics [14] and illustrate these definitions for our running example.

First, we instantiate Dung's abstract approach with arguments and attacks based on IGs.

Definition 15 (Argumentation framework). *Let $G_\mathcal{I} = (\mathbf{P}, \mathbf{A})$ be an IG, and let $\mathbf{E} \subseteq \mathbf{P}$ be an evidence set. An* argumentation framework (AF) defined by $G_\mathcal{I}$ and $\mathbf{E}$ *is a pair $(\mathcal{A}, \mathcal{C})$, where $\mathcal{A}$ is the set of all arguments on the basis of $G_\mathcal{I}$ and $\mathbf{E}$ as defined by Definition 9, and where $(A, B) \in \mathcal{C}$ iff $A, B \in \mathcal{A}$ and A attacks B (see Definition 11).*

An AF can be represented as a directed graph in which arguments are represented by circles and attacks are indicated by solid arcs ($\rightarrow$); an example of an AF is depicted in Figure 14.

Given an AF, we can use any semantics for AFs as defined in [14] for determining the dialectical status of arguments (cf. [21]). The theory of AFs is built around the notion of an *extension*, which is a set of arguments that is internally coherent and defends itself against attack.

Definition 16 (Dung extensions). *Let $(\mathcal{A}, \mathcal{C})$ be an AF defined by IG $G_\mathcal{I}$ and evidence set $\mathbf{E}$.*

- *A set of arguments $\mathcal{S} \subseteq \mathcal{A}$ is* conflict-free *if there do not exist $A, B \in \mathcal{S}$ such that $(A, B) \in \mathcal{C}$.*
- *An argument $A \in \mathcal{A}$ is* acceptable *with respect to some set of arguments $\mathcal{S} \subseteq \mathcal{A}$ iff for all arguments B such that $(B, A) \in \mathcal{C}$ there exists an argument $C \in \mathcal{S}$ such that $(C, B) \in \mathcal{C}$.*
- *A conflict-free set of arguments $\mathcal{S} \subseteq \mathcal{A}$ is an* admissible *extension iff every argument $A \in \mathcal{S}$ is acceptable with respect to $\mathcal{S}$.*
- *An admissible extension $\mathcal{S}$ is a* complete *extension iff $A \in \mathcal{S}$ whenever A is acceptable with respect to $\mathcal{S}$; $\mathcal{S}$ is the* grounded *extension iff $\mathcal{S}$ is the set inclusion minimal complete extension; $\mathcal{S}$ is a* preferred *extension iff $\mathcal{S}$ is a set inclusion maximal complete extension; and $\mathcal{S}$ is a* stable *extension iff it is preferred and $\forall B \notin \mathcal{S}, \exists A \in \mathcal{S}$ such that $(A, B) \in \mathcal{C}$.*

The acceptability of arguments in abstract argumentation frameworks can then be evaluated by establishing whether a given argument is a member of the various extensions. Arguments are then assigned a

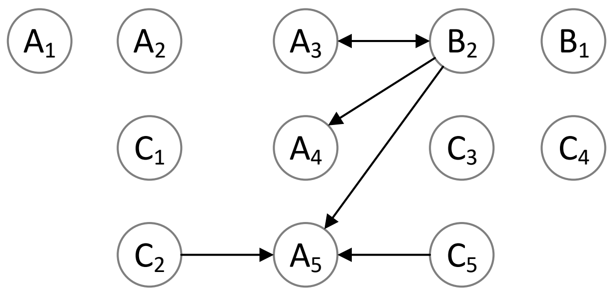


Fig. 14. AF corresponding to the IG of Figure 13.

dialectical status that can either be *'justified'*, *'overruled'*, or *'defensible'*, where informally an argument is justified if it survived the competition, overruled if it did not survive the competition, and defensible if it is involved in a tie.

Definition 17 (Justified, overruled and defensible arguments, adapted from [32])**.** *Let $(\mathcal{A}, \mathcal{C})$ be an argumentation framework.*

- *An argument is (i)* justified *under grounded semantics iff it is a member of the grounded extension, (ii)* overruled *under grounded semantics iff it is not justified under grounded semantics and it is attacked by an argument that is justified under grounded semantics, or (iii)* defensible *under grounded semantics iff it is neither justified nor overruled under grounded semantics.*
- *Let $T \in \{complete, preferred, stable\}$. An argument is (i)* justified *under T semantics iff it is a member of all T extensions, (ii)* overruled *under T semantics iff it is not a member of any T extension, or (iii)* defensible *under T semantics iff it is a member of some but not all T extensions.*

We now illustrate the evaluation of arguments based on IGs through our running example.

Example 39. *Consider the IG of Figure 13. To prevent this example from becoming too involved, we consider the following subset of arguments $\mathcal{A} = \{A_1, A_2, A_3, A_4, A_5, B_1, B_2, C_1, C_2, C_3, C_4, C_5\}$ and binary attack relation $\mathcal{C} = \{(A_3, B_2), (B_2, A_3), (B_2, A_4), (B_2, A_5), (C_2, A_5), (C_5, A_5)\}$ over $\mathcal{A}$ (see Examples 34, 36 and 37). The constructed AF $(\mathcal{A}, \mathcal{C})$ is visualised in Figure 14. The complete extensions of $(\mathcal{A}, \mathcal{C})$ are:*

$\mathcal{S}_1 = \{A_1, A_2, B_1, C_1, C_2, C_3, C_4, C_5\}$;
$\mathcal{S}_2 = \{A_1, A_2, B_1, B_2, C_1, C_2, C_3, C_4, C_5\}$
$\mathcal{S}_3 = \{A_1, A_2, A_3, A_4, B_1, C_1, C_2, C_3, C_4, C_5\}$.

Under complete semantics, $A_1, A_2, B_1, C_1, C_2, C_3, C_3, C_4, C_5$ are justified as they are members of all complete extensions, A_5 is overruled as it is attacked by a justified argument, and A_3, A_4 and B_2 are defensible. For the other semantics, the same statuses are assigned; for grounded semantics, this is the case as $\mathcal{S}_1$ is the set inclusion minimal complete extension. Furthermore, note that $\mathcal{S}_2$ and $\mathcal{S}_3$ are set inclusion maximal complete extensions for which it holds that $\forall B \notin \mathcal{S}_i, \exists A \in \mathcal{S}_i$ such that $(A, B) \in \mathcal{C}$ for $i = 2, 3$; hence, $\mathcal{S}_2$ and $\mathcal{S}_3$ are preferred and stable extensions. □

Dung's abstract argumentation approach has been extended with new elements, for instance by adding support relations to abstract argumentation frameworks (e.g. [11]) or by adding preference relations (e.g. so-called preference-based argumentation frameworks, or PAFs [1]), probabilities (see e.g. [16] for an overview), or weights [15] to AFs; a more complete overview is provided in [31]. We opt for the approach introduced by Dung for the evaluation of arguments as it is a well-studied and widely accepted approach in the field of computational argumentation. Moreover, the relations between Dung's fully abstract approach and formalisms for structured argumentation that are at an intermediate level of abstraction between concrete instantiating logics and Dung's approach, such as ASPIC$^+$ [21] and assumption-based argumentation (ABA) [7], have been previously investigated. In our IG-formalism, we have currently opted not to account for preferences, as these are typically not indicated in tools domain experts use. As the components of our argumentation formalism based on IGs are directly defined based on the elements that are accounted for in our IG-formalism, preferences are currently not accounted for in our argumentation formalism. As shown in work on structured argumentation with preferences (e.g. [21]), the structure of arguments is crucial in determining how preferences must be applied to attacks and one should be cautious in extending AFs with additional elements without taking the structure of arguments into account. There is some work on the relations between support relations in abstract argumentation frameworks and those at the inference level [30]. Relations between our proposed argumentation formalism and extended AFs such as [11] may be investigated in future research.

5.4. Satisfying rationality postulates

Caminada and Amgoud [10] studied rule-based argumentation systems and identified conditions under which unintuitive and undesirable results are obtained upon performing inference. They then defined principles, called rationality postulates, that can be used to judge the quality of a given rule-based argumentation system. More specifically, so-called consistency and closure postulates were formulated for systems allowing for strict and defeasible inferences. Since these postulates are widely accepted as important desiderata for structured argumentation formalisms, we prove in this section that these postulates are satisfied by instantiations of our argumentation formalism based on IGs.

5.4.1. Comparison of our argumentation formalism based on IGs to the ASPIC$^+$ framework

In proving satisfaction of [10]'s rationality postulates, we follow Modgil and Prakken [21], who proved satisfaction of these postulates for the ASPIC$^+$ framework. As noted earlier, in defining our argumentation formalism based on IGs we were inspired by the definitions of argument and attack as given in [21]. In Definition 9, we defined how arguments on the basis of an IG and an evidence set $\mathbf{E}$ are constructed. In step 2a of Definition 9, it is specified that an argument A with $\textsc{Conc}(A) = p$ can be constructed from arguments $A_1, \ldots, A_n$ if p is defeasibly deductively inferred from $\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)$ according to Definition 7 using a generalisation $g\colon \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)\} \rightarrow p$ in $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$. Hence, in terms of the terminology used in the ASPIC$^+$ framework, generalisations in $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$ can be interpreted as *domain-specific defeasible inference rules*[2] in ASPIC$^+$'s $\mathcal{R}_d$ that are applied when constructing arguments. Similarly, in step 2b of Definition 9 it is specified that an argument A with $\textsc{Conc}(A) = p$ can be constructed from $A_1, \ldots, A_n$ if p is strictly deductively inferred from $\textsc{Conc}(A_1)$, $\ldots, \textsc{Conc}(A_n)$ according to Definition 7 using a generalisation $g\colon \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n)\} \rightarrow p$ in $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$. Hence, generalisations in $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ can be interpreted as domain-specific *strict* inference rules in ASPIC$^+$'s $\mathcal{R}_s$. Finally in step 3 it is specified that an argument A with $\textsc{Conc}(A) = p$ can be constructed from an argument A' if p is abductively inferred from $\textsc{Conc}(A')$ according to Definition 8 using a $g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$, $g\colon \{p, p_1, \ldots, p_n\} \rightarrow \textsc{Conc}(A')$ for some propositions $p_1, \ldots, p_n \in \mathbf{P}$. Therefore, besides specifying aforementioned domain-specific defeasible and strict deduction rules, generalisations

[2]For details on using ASPIC$^+$ to model domain-specific defeasible and strict inference rules, the reader is referred to [22].

$g\colon \{q_1, \ldots, q_n\} \to q$ in $\mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{a}}$ also specify domain-specific *abduction* rules in $\mathcal{R}_d$, namely for every $i \in \{1, \ldots, n\}$ a rule can be specified in $\mathcal{R}_d$ that states that q_i can be defeasibly inferred from q.

Considering the different types of attacks that are defined in Section 5.2, rebuttal as defined in Section 5.2.1 is identical to rebuttal as defined for a special case of ASPIC$^+$, namely one in which conflict is based on the standard classical notion of negation. Undercutting as defined in Section 5.2.2 is a special case of undercutting as defined for ASPIC$^+$, as we only consider undercutters of inferences in case an exception is provided to a defeasible generalisation used in an inference step. Thus, of the types of attacks that are considered in our argumentation formalism, only alternative attack is not accounted for in ASPIC$^+$. Furthermore, in comparison to our argumentation formalism, Modgil and Prakken do not impose any additional restrictions on argument construction. Hence, to prove that instantiations of our argumentation formalism based on IGs satisfy rationality postulates, in Section 5.4.3 we focus on showing how alternative attack and the additional restrictions that are imposed on argument construction in our argumentation formalism can be taken account in the results and proofs provided in [21].

5.4.2. Additional definitions and assumptions

Following Modgil and Prakken [21], we introduce the following definitions. We define what it means for a set of propositions to be closed under strict generalisations.

Definition 18 (Closure under strict generalisations)**.** *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG and let $\mathbf{P}' \subseteq \mathbf{P}$. Then the* closure of $\mathbf{P}'$ under strict generalisations*, denoted* $\mathrm{CL}(\mathbf{P}')$*, is the smallest set containing $\mathbf{P}'$ and the consequent $Head(g)$ of any $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ whose antecedents $\mathbf{Tails}(g)$ are in $\mathrm{CL}(\mathbf{P}')$.*

Next, the terms directly consistent and indirectly consistent set are defined.

Definition 19 (Directly consistent set)**.** *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG and let $\mathbf{P}' \subseteq \mathbf{P}$. Then $\mathbf{P}'$ is* directly consistent *iff $\nexists p, q \in \mathbf{P}'$ such that $p = -q$.*

A set $\mathbf{P}'$ is indirectly consistent if its closure under strict generalisations is directly consistent.

Definition 20 (Indirectly consistent set)**.** *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG and let $\mathbf{P}' \subseteq \mathbf{P}$. Then $\mathbf{P}'$ is* indirectly consistent *iff* $\mathrm{CL}(\mathbf{P}')$ *is directly consistent.*

As noted by Caminada and Amgoud [10], one should search for ways to alter or constrain one's argumentation formalism in such a way that rationality postulates are satisfied. Accordingly, following Modgil and Prakken [21] we assume that IGs and evidence sets satisfy a number of properties. Similar to ASPIC$^+$, we leave the user free to make choices as to the strict and defeasible generalisations to include in $\mathbf{G} \subseteq \mathbf{A}$ and the observations to include in $\mathbf{E}$; however, some care needs to be taken in making these choices to ensure that the result of argumentation is guaranteed to be well-behaved. Specifically, to ensure rationality postulates are satisfied, we assume that evidence sets $\mathbf{E}$ are indirectly consistent (referred to as the *axiom consistency* assumption), and we assume that $\mathbf{G}$ is closed under transposition. Note that per definition every evidence set $\mathbf{E} \subseteq \mathbf{P}$ is a directly consistent set, as it is assumed in Definition 6 that for every $p \in \mathbf{E}$, $\neg p \notin \mathbf{E}$. Furthermore, all examples of IGs provided in this paper are axiom consistent, as they do not include generalisations $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ for which $\mathbf{Tails}(g) \subseteq \mathbf{E}$. Closure under transposition is one of the solutions proposed by Caminada and Amgoud to 'repair' an argumentation system to ensure rationality postulates are satisfied [10, p. 16], as it can help generate rules needed to obtain an intuitive outcome.

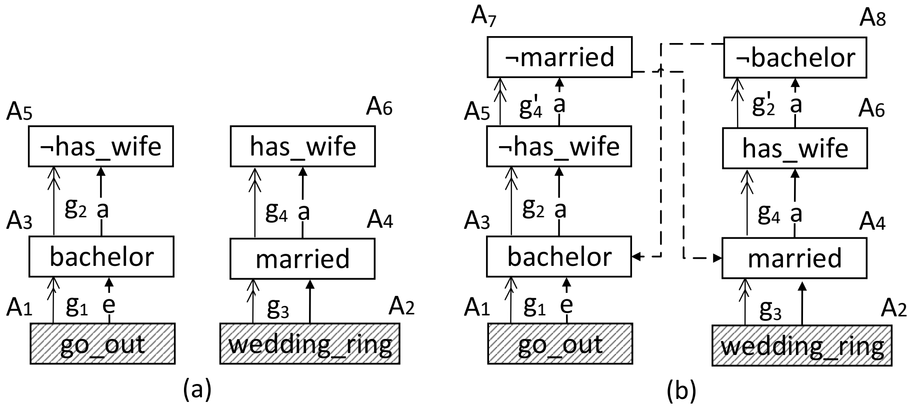


Fig. 15. Example of an IG for which $\mathbf{G}$ is not closed under transposition (a); adjustment to this IG, in which additional generalisations are included such that $\mathbf{G}$ is closed under transposition (b).

Definition 21 (Closure under transposition). *Let $G_{\mathcal{I}} = (\mathbf{P}, \mathbf{A})$ be an IG. A strict generalisation $g' \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ is a* transposition *of $g\colon \{p_1, \ldots, p_n\} \rightarrow p$ in $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$ iff g' is of the form $\{p_1, \ldots, p_{i-1}, -p, p_{i+1}, \ldots, p_n\} \rightarrow -p_i$ for some $1 \leqslant i \leqslant n$. We say that $\mathbf{G}$ is* closed under transposition *iff for all strict generalisations $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$, the transpositions of g are also in $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$.*

An AF $(\mathcal{A}, \mathcal{C})$ defined by an IG $G_{\mathcal{I}}$ that is axiom consistent and for which $\mathbf{G} \subseteq \mathbf{A}$ is closed under transposition is said to be *well defined*. In the remainder of this section, we assume that any given AF $(\mathcal{A}, \mathcal{C})$ is well defined. Note that most examples of IGs provided in this paper only include defeasible generalisations and not strict generalisations, and thus that AFs defined by these IGs are well defined. The following example, adapted from Caminada and Amgoud [10], illustrates closure under transposition and how ensuring it can help repair an argumentation system.

Example 40. *In the IG depicted in Figure 15a, strict abstractions $g_2\colon$* bachelor $\rightarrow$ ¬has_wife *and $g_4\colon$* married $\rightarrow$ has_wife *are included. $\mathbf{G}$ is not closed under transposition, as generalisations* has_wife $\rightarrow$ ¬bachelor *and* ¬has_wife $\rightarrow$ ¬married *are not included. Arguments A_5 and A_6 constructed on the basis of this IG have strict top inferences, as only step 2b of Definition 9 can be applied in constructing A_5 from A_3 and A_6 from A_4 using g_2 and g_4 in $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}}$, respectively. Note that, as* TOPINF(A_5) *and* TOPINF(A_6) *are strict, A_5 and A_6 are not involved in rebuttal. In fact, $\mathcal{C} = \emptyset$ for the AF corresponding to this IG, and hence under any semantics both A_5 and A_6 are justified. Thus, contradictory propositions* has_wife *and* ¬has_wife *are both justified at the same time, which is clearly undesirable and among other things violates the direct consistency postulate (see Theorem 1). In the IG depicted in Figure 15b, $\mathbf{G}$ is closed under transposition as additional generalisations* has_wife $\rightarrow$ ¬bachelor *and* ¬has_wife $\rightarrow$ ¬married *are now included. In the corresponding AF, A_7 directly rebuts A_4 and A_8 directly rebuts A_3 as* TOPINF(A_3) *and* TOPINF(A_4) *are defeasible. Then A_7 indirectly rebuts A_6 (on A_4) and A_8 indirectly rebuts A_5 (on A_3). Therefore, for this AF the more intuitive outcome is obtained that A_5 and A_6 cannot both be in the same extension at the same time.* □

Lastly, the following definitions introduce some terminology used in the below results. Following Modgil and Prakken [23], we define strict continuations in a slightly different way than in [21], but as noted by [23] this does not affect the proofs stated in [21].

Definition 22 (Strict continuations)**.** *Let* $(\mathcal{A}, \mathcal{C})$ *be an AF defined by IG* $G_{\mathcal{I}}$ *and evidence set* $\mathbf{E}$*. The set of* strict continuations *of a set of arguments from* $\mathcal{A}$ *is the smallest set satisfying the following conditions:*

(1) Any argument A *is a strict continuation of* $\{A\}$*.*
(2) If $A_1, \ldots, A_n$ *are arguments and* $S_1, \ldots, S_n$ *are sets of arguments such that for every* $i \in \{1, \ldots, n\}$*,* A_i *is a strict continuation of* S_i *and* $\{B_{n+1}, \ldots, B_m\}$ *is a (possibly empty) set of strict arguments, and* $g\colon \{\textsc{Conc}(A_1), \ldots, \textsc{Conc}(A_n), \textsc{Conc}(B_{n+1}), \ldots, \textsc{Conc}(B_m)\} \rightarrow p$ *is a strict generalisation in* $\mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$*, then argument* $A_1, \ldots, A_n, B_{n+1}, \ldots, B_m \rightharpoonup_g p$ *constructed from* $A_1, \ldots, A_n, B_{n+1}, \ldots, B_m$ *using* g *by applying step* 2b *of Definition 9 is a strict continuation of* $S_1 \cup \ldots \cup S_n$*.*

The maximal fallible sub-arguments of an argument B are those with the 'last' defeasible inferences in B. That is, they are the maximal sub-arguments of B on which B can be attacked.

Definition 23 (Maximal fallible sub-arguments)**.** *Let* $(\mathcal{A}, \mathcal{C})$ *be an AF defined by IG* $G_{\mathcal{I}}$ *and evidence set* $\mathbf{E}$*. The set* $M(B)$ *of the maximal fallible sub-arguments of* B *is defined such that for any* $B' \in \textsc{Sub}(B)$*, it holds that* $B' \in M(B)$ *iff:*

(1) $\textsc{TopInf}(B')$ *is defeasible, and;*
(2) There is no $B'' \in \textsc{Sub}(B)$ *such that* $B'' \neq B$*,* $B' \in \textsc{Sub}(B'')$ *and* B'' *satisfies condition 1.*

5.4.3. Proofs

We prove satisfaction of Caminada and Amgoud's consistency and closure postulates for complete semantics, which implies satisfaction of these postulates for grounded, preferred, and stable semantics. Caminada and Amgoud [10] also propose postulates for the intersection of extensions and their conclusion sets, but since their satisfaction directly follows from satisfaction of the postulates for individual extensions, these postulates will not be reconsidered.

First, a number of intermediate properties are proven. The intermediate result stated in Lemma 2 is identical to Lemma 37 of Modgil and Prakken [21], namely that any strict continuation B of a set of arguments $\{A_1, \ldots, A_n\}$ is acceptable with respect to $\mathcal{S}$ if all A_i are acceptable with respect to a set $\mathcal{S}$. The proof follows similar to Lemma 37 of [21], where alternative attack is now also considered.

Lemma 2. *Let* $(\mathcal{A}, \mathcal{C})$ *be an AF defined by IG* $G_{\mathcal{I}}$ *and evidence set* $\mathbf{E}$*. Let* $B \in \mathcal{A}$ *be a strict continuation of* $\{A_1, \ldots, A_n\}$*, and for* $i = 1, \ldots, n$*, let* A_i *be acceptable with respect to* $\mathcal{S} \subseteq \mathcal{A}$*. Then* B *is acceptable with respect to* $\mathcal{S}$*.*

Proof. Let A be any argument such that $(A, B) \in \mathcal{C}$. By Definition 11, A attacks B iff A rebuts B (on B'), A undercuts B (on B'), or A alternative attacks B (on B') for some $B' \in \textsc{Sub}(B)$ (see Definitions 12, 13, and 14). Here, it holds that $\textsc{TopInf}(B')$ is defeasible; more specifically:

(1) By Definition 12, A rebuts B (on B') iff B' is of the form $B''_1, \ldots, B''_n \twoheadrightarrow_g p$ for some $B''_1, \ldots, B''_n \in \mathcal{A}$, $p \in \mathbf{P}$ and hence iff $\textsc{TopInf}(B')$ is defeasible, and;
(2) By Definition 13, A undercuts B (on B') iff there exists an exception arc $x \in \mathbf{X}$ such that $x\colon \textsc{Conc}(A) \rightsquigarrow g$ and $\textsc{TopGen}(B') = g \in \mathbf{G}^{\mathrm{c}} \cup \mathbf{G}^{\mathrm{e}} \cup \mathbf{G}^{\mathrm{a}}_{\mathrm{d}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{d}}$. Hence, in constructing B' step 2b cannot be applied last, as this step can only be applied with strict generalisations $g \in \mathbf{G}^{\mathrm{a}}_{\mathrm{s}} \cup \mathbf{G}^{\mathrm{o}}_{\mathrm{s}}$. Therefore, step 2a of step 3 of Definition 9 is applied last in constructing B'. Thus, the last used inference in constructing B' is a defeasible deductive inference using $\textsc{TopGen}(B') = g$ (step 2a of Definition 9) or an abductive inference using $\textsc{TopGen}(B') = g$ (step 3 of Definition 9), and hence $\textsc{TopInf}(B')$ is defeasible, and;

(3) By Definition 14, A alternative attacks B (on B') iff $\text{TopInf}(B')$ is an abductive inference and hence iff $\text{TopInf}(B')$ is defeasible.

Hence, by definition of strict continuations (Definition 22), it must be that $(A,B) \in \mathcal{C}$ iff $(A,A_i) \in \mathcal{C}$ for some (possibly more than one) $A_i \in \{A_1,\dots,A_n\}$. Specifically, if A does not undercut, rebut or alternative attack some A_i, then this contradicts that $(A,B) \in \mathcal{C}$. Thus, we have shown that if $(A,B) \in \mathcal{C}$, then $(A,A_i) \in \mathcal{C}$ for some $A_i \in \{A_1,\dots,A_n\}$. By assumption, A_i is acceptable with respect to $\mathcal{S}$, thus $\exists C \in \mathcal{S}$ such that $(C,A) \in \mathcal{C}$. Thus, B is acceptable with respect to $\mathcal{S}$. ■

The intermediate result stated in Lemma 3 is similar to Proposition 8 of Modgil and Prakken [21]. Compared to Proposition 8 of [21], in which no assumptions are made regarding A, we now assume that A is defeasible with a strict top inference or that A is strict, as these are the only cases needed in our proof of Theorem 1. As Modgil and Prakken do not impose any restrictions on argument construction in their formalism, a result proven by Caminada and Amgoud [10] (i.e. Lemma 6 of [10]) can be directly used to complete their proof. Below, we show that the restrictions that are imposed on argument construction in our argumentation formalism based on IGs do not restrict the construction of strict continuations, and hence that the proof can similarly be completed.

Lemma 3. *Let $(\mathcal{A},\mathcal{C})$ be an AF defined by IG $G_\mathcal{I}$ and evidence set* $\mathbf{E}$. *Let A and B be arguments in $\mathcal{A}$ such that B is defeasible,* $\text{Conc}(A) = -\text{Conc}(B)$. *Let A be strict or let A be defeasible with* $\text{TopInf}(A)$ *strict. Then for all $B' \in M(B)$, there exists a strict continuation A^+ of $(M(B) \setminus \{B'\}) \cup \{A\}$ such that A^+ rebuts B on B'.*

Proof. Let A be strict or let A be defeasible with $\text{TopInf}(A)$ strict. Let B be defeasible with $\text{Conc}(A) = -\text{Conc}(B)$. First, note that according to Definition 22 any strict continuation of a given set of arguments from $\mathcal{A}$ is either (1) A if the set of arguments under consideration is $\{A\}$ (Definition 22, condition 1), or (2) is constructed by applying step 2b of Definition 9 one or more (but finitely many) times (Definition 22, condition 2). As restrictions are imposed on argument construction in our argumentation formalism based on IGs, we first show that in constructing any strict continuation A^+ of $(M(B) \setminus \{B'\}) \cup \{A\}$ step 2b of Definition 9 can be applied without restrictions.

Generally, in applying step 2b of Definition 9 an argument C with $\text{Conc}(C) = p$ is constructed from arguments $C_1,\dots,C_n$ by strictly deductively inferring p from propositions $\text{Conc}(C_1),\dots,\text{Conc}(C_n)$ according to Definition 7 using a generalisation $g\colon \text{Conc}(C_1),\dots,\text{Conc}(C_n) \to p$ in $\mathbf{G}^\text{a}_\text{s} \cup \mathbf{G}^\text{o}_\text{s}$. In Definition 7 no constraints are imposed on performing deduction with strict generalisations $g \in \mathbf{G}^\text{a}_\text{s} \cup \mathbf{G}^\text{o}_\text{s}$; in particular, the only constraint that is imposed is in condition 2 of this definition, where constraints are imposed on performing deduction with *defeasible* generalisations in $\mathbf{G}^\text{e}$ (i.e. Pearl's constraint). The only other case in which step 2b of Definition 9 cannot be applied in constructing an argument C using a $g \in \mathbf{G}^\text{a}_\text{s} \cup \mathbf{G}^\text{o}_\text{s}$ is in case the same g was already used in the previous construction step to construct an argument $C' \in \text{ImmSub}(C)$, namely by applying step 3 of Definition 9. Now again consider argument A. By assumption, A is strict or $\text{TopInf}(A)$ strict, and therefore step 3 of Definition 9, which specifies a defeasible inference, could not have been applied last in constructing A; therefore, no restrictions are imposed on constructing strict continuations A^+ of $(M(B) \setminus \{B'\}) \cup \{A\}$ in our argumentation formalism. By assumption, $(\mathcal{A},\mathcal{C})$ is well defined and, therefore, closed under transposition; hence, by straightforward generalisation of Lemma 6 in [10] one can construct a strict continuation A^+ that continues $(M(B) \setminus \{B'\}) \cup \{A\}$ with strict inferences and that concludes $-\text{Conc}(B')$. Since by construction of $M(B)$, B' has a defeasible top inference and therefore A^+ rebuts B'. But then A^+ also rebuts B. ■

The intermediate result stated in Lemma 4 is identical to Lemma 38 of [21].

Lemma 4. *Let $(\mathcal{A},\mathcal{C})$ be an AF defined by IG $G_{\mathcal{I}}$ and evidence set* $\mathbf{E}$. *Let $A \in \mathcal{A}$ be acceptable w.r.t. admissible extension $\mathcal{S} \subseteq \mathcal{A}$. Let $\mathcal{S}' = \mathcal{S} \cup \{A\}$. Then $\forall B \in \mathcal{S}'$, neither $(A,B) \in \mathcal{C}$ nor $(B,A) \in \mathcal{C}$.*

Proof. Suppose for contradiction that: (1) $\exists B \in \mathcal{S}'$ such that $(A,B) \in \mathcal{C}$. As $B \in \mathcal{S}'$, it follows that B is acceptable w.r.t. $\mathcal{S}$, as either $B = A$, which is acceptable w.r.t. $\mathcal{S}$ by assumption, or B is an element of admissible extension $\mathcal{S}$. Hence $\exists C \in \mathcal{S}$ such that $(C,A) \in \mathcal{C}$. Then, as A is acceptable w.r.t. $\mathcal{S}$, $\exists D \in \mathcal{S}$ such that $(D,C) \in \mathcal{C}$, contradicting $\mathcal{S}$ is conflict-free; (2) $\exists B \in \mathcal{S}'$ such that $(B,A) \in \mathcal{C}$. As A is acceptable w.r.t. $\mathcal{S}$, $\exists C \in \mathcal{S}$ such that $(C,B) \in \mathcal{C}$, contradicting $\mathcal{S}$ is conflict-free. ■

The result stated in Lemma 5 is identical to Lemma 35-2 of Modgil and Prakken [21], namely that an argument A attacks an argument B iff A attacks some sub-argument B' of B. Compared to Lemma 35-2 of [21], alternative attack is now also considered in the proof.

Lemma 5. *Let $(\mathcal{A},\mathcal{C})$ be an AF defined by IG $G_{\mathcal{I}}$ and evidence set* $\mathbf{E}$. *Let $A, B \in \mathcal{A}$. Then $(A,B) \in \mathcal{C}$ iff $(A,B') \in \mathcal{C}$ for some $B' \in \textsc{Sub}(B)$.*

Proof. By Definition 11, $(A,B) \in \mathcal{C}$ iff A rebuts B (on B'), A undercuts B (on B'), or A alternative attacks B (on B') for some $B' \in \textsc{Sub}(B)$ (see Definitions 12, 13, and 14); hence, also $(A,B') \in \mathcal{C}$. ■

The intermediate result stated in Lemma 6 is identical to Proposition 10 of [21].

Lemma 6. *Let $(\mathcal{A},\mathcal{C})$ be an AF defined by IG $G_{\mathcal{I}}$ and evidence set* $\mathbf{E}$. *Let $A \in \mathcal{A}$ be acceptable with respect to admissible extension $\mathcal{S} \subseteq \mathcal{A}$. Then $\mathcal{S}' = \mathcal{S} \cup \{A\}$ is conflict-free.*

Proof. We need to show that there do not exist $B, C \in \mathcal{S}'$ such that $(B,C) \in \mathcal{C}$. As $\mathcal{S}$ is an admissible extension, $\mathcal{S}$ is conflict free: hence, there do not exist $B, C \in \mathcal{S}$ such that $(B,C) \in \mathcal{C}$. Thus, we need to show that $(A,A) \notin \mathcal{C}$, and neither $(A,B) \in \mathcal{C}$ nor $(B,A) \in \mathcal{C}$ for all $B \in \mathcal{S}$. As by assumption A is acceptable with respect to $\mathcal{S}$, this follows directly from Lemma 4. ■

Theorem 1, corresponding to the direct consistency postulate, states that the conclusions of arguments in an admissible extension (and so by implication in a complete extension) are directly consistent. The conclusions of arguments in an extension should not be contradictory, as this leads to what Caminada and Amgoud call 'absurdities' [10, p. 15] in that two contradictory statements can then be justified at the same time.

Theorem 1 (Direct consistency)**.** *Let $(\mathcal{A},\mathcal{C})$ be an AF defined by IG $G_{\mathcal{I}}$ and evidence set* $\mathbf{E}$. *Then for all admissible extensions $\mathcal{S}$ of AF it holds that the set $\{\textsc{Conc}(A) \mid A \in \mathcal{S}\}$ is directly consistent.*

Proof. Let $\mathcal{S}$ be an admissible extension of AF and let A and B be arguments in $\mathcal{S}$. We show that if $\textsc{Conc}(A) = q$, $\textsc{Conc}(B) = r$ with $q = -r$ (i.e. $\{\textsc{Conc}(A) \mid A \in \mathcal{S}\}$ is not directly consistent), then this leads to a contradiction:

(1) If A is a strict argument, and:
  1.1 if B is also strict, then this contradicts our *axiom consistency* assumption on evidence sets $\mathbf{E}$;
  1.2 if B is a defeasible argument, and:

1.2.1 if B has a defeasible top inference, then A rebuts B (on B) by Definition 12, as a negation arc $n\colon \textsc{Conc}(A) \leftrightsquigarrow \textsc{Conc}(B)$ exists in $\mathbf{N}$ (as $q = -r$). Hence, this contradicts that $\mathcal{S}$ is conflict-free.

1.2.2 if B has a strict top inference, then by Lemma 3 there exists a strict continuation A^+ of $(M(B)\setminus\{B'\})\cup\{A\}$ for every $B' \in M(B)$ such that A^+ rebuts B on B'; hence, $(A^+, B) \in \mathcal{C}$. By our Lemma 2, A^+ is acceptable with respect to $\mathcal{S}$, and by Lemma 6, $\mathcal{S} \cup \{A^+\}$ is conflict-free, contradicting that $(A^+, B) \in \mathcal{C}$.

(2) If A is a defeasible argument and B is a strict argument, then the result follows similar to case 1.2 with the roles of arguments A and B reversed.

(3) If A and B are defeasible arguments, and:

3.1 if $\textsc{TopInf}(A)$ or $\textsc{TopInf}(B)$ is defeasible, then the result follows similar to case 1.2.1 (either with the roles of arguments A and B as they currently are or with their roles reversed).

3.2 if $\textsc{TopInf}(A)$ and $\textsc{TopInf}(B)$ are strict, then the result follows similar to case 1.2.2. ■

The result stated in Lemma 7 is identical to Lemma 35-3 of [21].

Lemma 7. *Let $(\mathcal{A}, \mathcal{C})$ be an AF defined by IG $G_\mathcal{I}$ and evidence set* $\mathbf{E}$. *Let $\mathcal{S} \subseteq \mathcal{A}$ and let $A \in \mathcal{S}$ with $A' \in \textsc{Sub}(A)$. Then A' is acceptable with respect to $\mathcal{S}$ if A is acceptable with respect to $\mathcal{S}$.*

Proof. Assume that A is acceptable with respect to $\mathcal{S}$. We need to prove that for every argument B such that $(B, A') \in \mathcal{C}$, $\exists C \in \mathcal{S}$ such that $(C, B) \in \mathcal{C}$. Let $B \in \mathcal{A}$ and assume that $(B, A') \in \mathcal{C}$. By Lemma 5, $(B, A) \in \mathcal{C}$. Then, as A is acceptable with respect to $\mathcal{S}$, $\exists C \in \mathcal{S}$ such that $(C, B) \in \mathcal{C}$. Hence, A' is acceptable with respect to $\mathcal{S}$. ■

Below, Caminada and Amgoud's [10] closure and indirect consistency postulates are stated. Informally, the closure postulates state that the conclusions returned by an argumentation system should be 'complete' [10, p. 16]. The sub-argument closure postulate states that for any argument A in a complete extension $\mathcal{S}$, all sub-arguments of A are also in $\mathcal{S}$.

Theorem 2 (Sub-argument closure)**.** *Let $(\mathcal{A}, \mathcal{C})$ be an AF defined by IG $G_\mathcal{I}$ and evidence set* $\mathbf{E}$. *Then for all complete extensions $\mathcal{S}$ of AF it holds that if an argument A is in $\mathcal{S}$ then all sub-arguments $A' \in \textsc{Sub}(A)$ of A are in $\mathcal{S}$.*

Proof. Let $\mathcal{S}$ be a complete extension of AF, let $A \in \mathcal{S}$ and let $A' \in \textsc{Sub}(A)$. Then A' is acceptable with respect to $\mathcal{S}$ by Lemma 7. Then $\mathcal{S} \cup \{A'\}$ is conflict-free by Lemma 6. Hence, since $\mathcal{S}$ is complete, it holds that $A' \in \mathcal{S}$. ■

Theorem 3, corresponding to the strict closure postulate, states that the conclusions of arguments in a complete extension are closed under strict inference.

Theorem 3 (Closure under strict inferences)**.** *Let $(\mathcal{A}, \mathcal{C})$ be an AF defined by IG $G_\mathcal{I}$ and evidence set* $\mathbf{E}$. *Let $\mathcal{S}$ be a complete extension of AF. Then* $\{\textsc{Conc}(A) \mid A \in \mathcal{S}\} = \mathrm{CL}(\{\textsc{Conc}(A) \mid A \in \mathcal{S}\})$.

Proof. It suffices to show that any strict continuation X of $\{A \mid A \in \mathcal{S}\}$ is in $\mathcal{S}$. By Lemma 2, any such X is acceptable with respect to $\mathcal{S}$. By Lemma 6, $\mathcal{S} \cup \{X\}$ is conflict-free. Hence, since $\mathcal{S}$ is complete, it follows that $X \in \mathcal{S}$. ■

Finally, Theorem 4, corresponding to the indirect consistency postulate, states the mutual consistency of the strict closure of conclusions of arguments in a complete extension.

Theorem 4 (Indirect consistency)**.** *Let* $(\mathcal{A}, \mathcal{C})$ *be an AF defined by IG* $G_{\mathcal{I}}$ *and evidence set* $\mathbf{E}$. *Let* $\mathcal{S}$ *be a complete extension of AF. Then* $\{\textsc{Conc}(A) \mid A \in \mathcal{S}\}$ *is indirectly consistent.*

Proof. The result follows from Theorems 1 and 3. ■

To conclude this section, we have shown that instantiations of our argumentation formalism based on IGs satisfy Caminada and Amgoud's [10] consistency and closure postulates. Satisfaction of these postulates warrants the sound definition of instantiations of our argumentation system and implies that anomalous results as identified by [10] are avoided.

6. Related work

In this paper, we have proposed an argumentation formalism based on IGs that allows for both deductive and abductive argumentation and which instantiates Dung's [14] abstract approach. Earlier work by Bex [4, 5] is related, although only his integrated theory [5] is purely argumentation-based; the relation to [5] was discussed in the introduction. The hybrid theory proposed by Bex [4] is a formal account of reasoning about evidence in which deduction and abduction are used in constructing evidential arguments and causal stories, which are completely separate entities with their own definitions related to conflict and evaluation. In comparison, our argumentation formalism based on IGs allows for the construction of both deductive and abductive arguments. Moreover, Bex's hybrid theory does not allow for most types of mixed inference with causal and evidential generalisations and abstractions, and largely avoids the problems associated with mixed inference as identified by Pearl [26] and as identified in the current paper. Bench-Capon and Prakken [3] offer a formalisation of Aristotle's practical syllogism within a logic for defeasible argumentation that is essentially a preliminary version of ASPIC$^+$ [21]. This approach allows for reasoning about alternative goals and values to justify actions, which is akin to performing abductive inference. In formalising this syllogism, Bench-Capon and Prakken only consider the abductive nature of reasoning about desires on the basis of beliefs and goals, whereas we offer a general account of abductive (and deductive) argumentation. Booth and colleagues [8] propose a top-down approach by developing a model of abduction in abstract argumentation [14] and instantiating their approach with abductive logic programs [20]. In comparison to our bottom-up approach, their approach does not allow for mixed abductive-deductive inference with different types of information.

The argumentation formalism presented in this paper is based on a version of the graph-based IG-formalism that considers causal, evidential, abstraction, and other types of generalisations, as well as generalisations that include enabling conditions. Most related formalisms for inference with these types of information are logic-based [4, 5, 13, 18, 25, 29, 34, 35] and do not consider the constraints on performing inference that need to be imposed. Poole's Theorist framework [29] and Shanahan's approach [34] only allow for causal defaults; complications with reasoning using both causal and evidential defaults as identified by Pearl [26] are thus avoided. The approaches of Ortiz Jr. [25] and Shoham [35] similarly only allow for inference with causal rules, but in contrast to [29, 34] also include enabling conditions. The formal logical model of abductive reasoning proposed by Josephson and Josephson [18] allows for explaining observations using causal rules. The approach by Console and Dupré [13] is similar in nature to [18] but also allows for abduction using abstractions, as discussed in Section 2.2.

Graph-based formalisms for reasoning with causality information have also been proposed, notably Pearl's causal diagrams [27]. Pearl provides a framework for causal inference in which diagrams are queried to determine if the assumptions available are sufficient for identifying causal effects. Compared to our IG-formalism and our argumentation formalism based on IGs, this framework does not allow for capturing asymmetric conflicts such as exceptions in the graph. Moreover, causal diagrams require probabilistic quantification to be queried, while IGs are qualitative.

7. Conclusion

In this paper, we have proposed an argumentation formalism that allows for both deductive and abductive argumentation, the latter of which has received relatively little attention in argumentation. Our argumentation formalism is based on an extended version of our previously proposed IG-formalism [39], where in addition to causal and evidential generalisations we now also allow for abstractions and other types of generalisations, thereby increasing the expressivity of our IG-formalism. We have identified conditions under which performing inference with abstractions can lead to undesirable results, thereby extending the set of inference constraints imposed by Pearl's C-E system for reasoning with causal and evidential information [26]. Moreover, we have identified exceptional circumstances under which the constraints of Pearl's C-E system should not be imposed, namely in case enabling conditions are provided under which a generalisation may be used in performing inference. Based on these constraints and our conceptional analysis of reasoning about evidence, we have defined how deduction and abduction may be performed with IGs. We have then formally proven that arguments constructed in our argumentation formalism based on IGs indeed adhere to these constraints. In the paper, we have focused on the constraints that need to be imposed on performing inference with pairs of generalisations, which cover Pearl's original constraints and local constraints on performing inference with abstractions. In future work, additional inference constraints may be imposed for longer chains of inferences involving more specific combinations of generalisations, granted that the total set of constraints is consistent. Furthermore, as causality is a contentious topic, our argumentation formalism may be extended in future work by allowing for meta-argumentation about labels of generalisations, as well as other elements of IGs.

Besides allowing for rebuttal and undercutting attack, which are among the types of attacks that are typically distinguished in structured argumentation [21, 28], we have also defined the notion of alternative attack among arguments based on IGs, a concept based on the notion of competing alternative explanations that is inspired by [3, 5]. Alternative attack captures a crucial aspect of abductive reasoning, namely that of conflict between abductively inferred conclusions [18]. We have contributed to the literature on computational argumentation by allowing for the formal evaluation of arguments involved in this type of conflict. Moreover, we have shown that instantiations of our argumentation formalism satisfy key rationality postulates [10], which warrants the sound definition of instantiations of our argumentation system and implies that anomalous results such as issues regarding inconsistency and non-closure as identified by [10] are avoided.

Our argumentation formalism generates an abstract AF as in Dung [14] and thus allows arguments to be formally evaluated according to Dung's argumentation semantics. By formalising analyses performed by domain experts using the informal reasoning tools they are familiar with (e.g. mind maps) as IGs as an intermediary step, this allows for the evaluation of IGs using computational argumentation, as well as using other formal systems such as BNs [39].

References

[1] L. Amgoud and C. Cayrol, A model of reasoning based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34: 197–215, 2002.

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

[3] T.J.M. Bench-Capon and H. Prakken. Justifying actions by accruing arguments. In P.E. Dunne and T.J.M. Bench-Capon, eds., *Computational Models of Arguments: Proceedings of COMMA 2006*, vol. 144, pp. 247–258. IOS Press, 2006.

[4] F. Bex. *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*. Springer, 2011.
[5] F. Bex. An integrated theory of causal stories and evidential arguments. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law*, pp. 13–22. ACM Press, 2015.
[6] F. Bex, H. Prakken, C.A. Reed, and D. Walton. Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law*, 11(2–3): 125–165, 2003.
[7] A. Bondarenko, P. Dung, R. Kowalski, and F. Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93: 63–101, 1997.
[8] R. Booth, D. Gabbay, S. Kaci, T. Rienstra, and L. van der Torre. Abduction and dialogical proof in argumentation and logic programming. In T. Schaub, G. Friedrich, and B. O'Sullivan, eds., *Proceedings of the Twenty-First European Conference on Artificial Intelligence*, vol. 263, pp. 117–122. IOS Press, 2014.
[9] S.W. van den Braak, H. van Oostendorp, H. Prakken, and G.A.W. Vreeswijk. Representing narrative and testimonial knowledge in sense-making software for crime analysis. In E. Francesconi, G. Sartor, and D. Tiscornia, eds., *Legal Knowledge and Information Systems: JURIX 2008: The Twenty-First Annual Conference*, vol. 189, pp. 160–169. IOS Press, 2008.
[10] M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5–6): 286–310, 2007.
[11] C. Cayrol and M.-C. Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation. In L. Godo, ed., *Proceedings of the Eight European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, vol. 3571, pp. 378–389. Springer, 2005.
[12] P.W. Cheng and L.R. Novick. Causes versus enabling conditions. *Cognition*. 40: 83–120, 1990.
[13] L. Console and D.T. Dupré. Abductive reasoning with abstraction axioms. In G. Lakemeyer and B. Nebel, eds., *Foundations of Knowledge Representation and Reasoning*, vol. 810, pp. 98–112. Springer, 1994.
[14] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2): 321–357, 1995.
[15] P.E. Dunne, A. Hunter, P. McBurney, S. Parsons, and M. Wooldridge. Weighted argument systems: basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175: 457–486, 2011.
[16] A. Hunter and M. Thimm. Probabilistic reasoning with abstract argumentation frameworks. *Journal of Artificial Intelligence Research*, 59: 565–611, 2017.
[17] F.V. Jensen and T.D. Nielsen. *Bayesian Networks and Decision Graphs*. Springer, 2nd ed., 2007.
[18] J.R. Josephson and S.G. Josephson. *Abductive Inference: Computation, Philosophy, Technology*. Cambridge University Press, 1994.
[19] J.B. Kadane and D.A. Schum. *A Probabilistic Analysis of the Sacco and Vanzetti Evidence.* John Wiley & Sons Inc., 1996.
[20] A.C. Kakas, R. Kowalski, and F. Toni. Abductive logic programming. *Journal of Logic and Computation*, 2(6): 719–770, 1993.
[21] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195, 361–397, 2013.
[22] S. Modgil and H. Prakken. The ASPIC^{+} framework for structured argumentation: a tutorial. *Argument and Computation*, 5(1): pp. 31–62, 2014.
[23] S. Modgil and H. Prakken. Abstract rule-based argumentation. In P. Baroni, D. Gabbay, M. Giacomin, and L. van der Torre, eds., *Handbook of Formal Argumentation*, pp. 286-361. College Publications, 2018.
[24] A. Okada, S.J. Buckingham Shum, and T. Sherborne, eds., *Knowledge Cartography: Software Tools and Mapping Techniques.* Springer, 2nd ed., 2014.
[25] C.L. Ortiz Jr. A commonsense language for reasoning about causation and rational action. *Artificial Intelligence*, 111(1-2), 73–130, 1999.
[26] J. Pearl. Embracing causality in default reasoning. *Artificial Intelligence*, 35(2): 259–271, 1988.
[27] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2nd ed., 2009.
[28] J. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person*. MIT Press, 1995.
[29] D. Poole. Representing diagnosis knowledge. *Annals of Mathematics and Artificial Intelligence*, 11(1–4):33–50, 1994.
[30] H. Prakken. On support relations in abstract argumentation as abstractions of inferential relations. In T. Schaub, G. Friedrich, and B. O'Sullivan, eds., *Proceedings of the Twenty-First European Conference on Artificial Intelligence*, vol. 263, pp. 735–740. IOS Press, 2014.
[31] H. Prakken. Historical overview of formal argumentation. In P. Baroni, D. Gabbay, M. Giacomin, and L. van der Torre, eds., *Handbook of Formal Argumentation*, pp. 73-141. College Publications, 2018.
[32] H. Prakken and G. Vreeswijk. Logics for defeasible argumentation. In R. Goebel and F. Guenthner, *Handbook of Philosophical Logic*, vol. 4, pp. 219–318. Springer, 2002.
[33] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1–2): 81–132, 1980.
[34] M. Shanahan. Prediction is deduction but explanation is abduction. In N.S. Sridharan, ed., *Proceedings of International Joint Conference on Artificial Intelligence 89*, 1055–1060. Morgan Kaufmann, 1989.

[35] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*, MIT Press, 1988.

[36] N. Timmers. The hybrid theory in practice: a case study at the Dutch police force. Master's thesis, Utrecht University, The Netherlands, 2017.

[37] R. Wieten, F. Bex, H. Prakken, and S. Renooij. Exploiting causality in constructing Bayesian networks from legal arguments. In M. Palmirani, ed., *Legal Knowledge and Information Systems. JURIX 2018: The Thirty-First Annual Conference*, vol. 313, pp. 151–160. IOS Press, 2018.

[38] R. Wieten, F. Bex, H. Prakken, and S. Renooij. Deductive and abductive reasoning with causal and evidential information. In H. Prakken, S. Bistarelli, F. Santini, and C. Taticchi, eds., *Computational Models of Argument. Proceedings of COMMA 2020*, vol. 326, pp. 383–394. IOS Press, 2020.

[39] R. Wieten, F. Bex, H. Prakken, and S. Renooij. Information graphs and their use for Bayesian network construction. Manuscript submitted to the International Journal of Approximate Reasoning, 2020.

[40] J.H. Wigmore. *The Principles of Judicial Proof.* Little, Brown and Company, 1913.