Informal Logic

INFORMAL LOGIC

Informalizing Formal Logic

Antonis Kakas

Explore this journal

Cite this article

Article abstract

This paper presents a way in which formal logic can be understood and
reformulated in terms of argumentation that can help us unify formal and
informal reasoning. Classical deductive reasoning will be expressed entirely in
terms of notions and concepts from argumentation so that formal logical
entailment is equivalently captured via the arguments that win between those
supporting concluding formulae and arguments supporting contradictory
formulae. This allows us to go beyond Classical Logic and smoothly connect it
with human reasoning, thus providing a uniform argumentation-based view of
both informal and formal logic.

Informalizing Formal Logic

ANTONIS C. KAKAS

Department of Computer Science
University of Cyprus
1 University Avenue
2109 Nicosia, Cyprus
antonis@ucy.ac.cy

Abstract: This paper presents a way in which formal logic can be understood and reformulated in terms of argumentation that can help us unify formal and informal reasoning. Classical deductive reasoning will be expressed entirely in terms of notions and concepts from argumentation so that formal logical entailment is equivalently captured via the arguments that win between those supporting concluding formulae and arguments supporting contradictory formulae. This allows us to go beyond Classical Logic and smoothly connect it with human reasoning, thus providing a uniform argumentation-based view of both informal and formal logic.

Résumé: Cet article montre comment la logique formelle peut être comprise et reformulée en termes d'argumentation de manière à nous aider à unifier le raisonnement formel et non formel. Le raisonnement déductif classique sera entièrement exprimé en termes de notions et de concepts issus de l'argumentation, de sorte que l'implication logique formelle soit saisie de manière équivalente via les arguments gagnants entre ceux qui soutiennent les formules de conclusion et ceux qui soutiennent des formules contradictoires. Cela nous permet d'aller au-delà de la logique classique et de la relier facilement au raisonnement humain, fournissant ainsi une vision uniforme, basée sur une argumentation, de la logique non formelle et formelle.

1. Introduction

Informal logic (Groarke 2017; Johnson and Blair 2000; Walton 2008) is generally regarded as a completely different form of reasoning to that of formal logic used in mathematics and science. Its main characteristic, with which it is often equated, is that of argumentation and in particular the study of "real-life everyday" argument so as to distinguish its concern as outside of any formal setting.

Yet, informal and formal logic share a common root, namely that of understanding human reasoning, and it is natural to ask how they might be related. One direction of linking the two forms of logical reasoning would be to attempt to give a precise or formal definition of informal logic. This can be a difficult task exactly because of the all-encompassing nature of informal logic, which widely covers human reasoning, and its emphasis on the practical aspects of this resulting from the emersion of informal reasoning in real-life settings. There are several works aiming to capture informal logic in a precise formal setting such as that found in the article "Formalizing informal logic" (Walton and Gordon 2015) where informal logic is placed in the formal argumentation framework setting of the Carneades argumentation system (Gordon and Walton 2006).

This paper is concerned with the question in the other direction: how can formal logic be captured within the framework of informal logic and argumentation? We are interested to understand in which way the same "apparatus" of informal logic and specifically that of argumentation, can be instantiated to give formal logical entailment. In other words, we are interested to understand if and how informal logic can constitute an overarching framework that encompasses the full range of logical reasoning?

Links between informal logic and formal logic are well known, e.g. argument schemes in informal logic are related to proof rules in formal logic and formal derivations are connected to the dialectic process of dialogues, games and argumentation. For example, dialogical logic (Lorenz 2001) is a logical framework that builds on such a synthesis of concepts from informal and formal logic. Our approach falls within this effort of bridging informal and formal logic. It is based on the relatively recent work on argumentation developed within the field of artificial intelligence in an effort to formalize and build systems that exhibit natural or human like intelligence (see Baroni et al. 2018; Bench-Capon and Dunne 2007; Besnard and Hunter 2008; and references within these). This computational approach to argumentation stems primarily from the study of non-monotonic reasoning in AI, but it is also rooted in the tradition of informal logic through philosophical studies of logic and argumentation, such as the works of Pollock (1987) and Toulmin (1958) and the chapter, "Argumentation and Artificial Intelligence",

in the *Handbook of Argumentation Theory* (van Eemeren et al. 2014).

The framework of argumentation in AI contains two main parts. One part, referred to as 'abstract argumentation', is concerned with the abstract formal framework of argumentation and the semantic acceptability of arguments. The second part, referred to as 'structured argumentation', studies representations of the abstract framework within some application language, as for example, the language of some logical framework. They have been used to develop computational models for several problems that can be categorized as problems within the realm of natural intelligence and informal logic. These problems include default reasoning and commonsense explanations, autonomous decision making in multi-agent environments, reasoning about actions and change and narrative comprehension, dialogues and debate analysis and many others (see the book [Rahwan and Simari 2009] for examples of studies of these applications). The multitude and variety of these problems testify to the appropriateness of the framework of argumentation in AI for the study of informal logic.

The purpose of this paper is to present a way that formal logic can be linked to informal logic, by showing how formal logical reasoning can be understood as a specific case of this AI argumentation framework. We will present a way that we can reconstruct formal logical reasoning entirely in terms of argumentation thus enabling us to view this as a process of dialectic argumentation. Informalizing formal logic will be recognized as a limiting case of dialectic argumentation where the arena of arguments together with the notions of counter-argument and defending argument are all rigidly fixed. In this limiting case most of the flexibility of (informal) argumentation, such as that of building arguments from loose associations or that of choosing which arguments to use in the argumentation process, is frozen by making particular and absolute choices for the various parameters that shape an argumentation framework.

The rigidity of the argumentation framework for formal logic is to be expected since our task is to precisely recover strict formal reasoning. The importance, though, of this reformulation of formal logical reasoning is that the strictness in the argumentation framework can be subsequently relaxed in cases where this is appropriate. For

example, such cases include problems where the given premises are inconsistent or when the premises contain loose associations such as defaults of commonsense knowledge and/or other information relating to human biases when we are in the realm of general human reasoning. Argumentation will thus provide a uniform way of capturing both formal and informal reasoning, smoothly moving from one to the other by varying the degree of flexibility within the argumentation framework.

There are several technical considerations that play an important role in realizing this uniform link of formal logic to argumentation. One such consideration concerns the formal reasoning process of proof by contradiction whose inclusion within argumentation is done more appropriately, not as an argument scheme or proof rule, as is typical in formal proof systems, but rather as part of the dialectic process of evaluation of arguments and in particular the semantic identification of self-defeating arguments. This enables a simple and direct connection of formal logic to argumentation. Another important technical element concerns the formal notion of logical derivation of a concluding formula. This will be understood in a wider sense, where a dialectic process of acceptability of the formulae is accompanied by an analogous but dual dialectic process for the non-acceptability of the negation of the formula. In other words, a logical conclusion cannot, in general, be derived through a single dialectic derivation of this formula. And although when the premises of a given logical theory are consistent, this is possible, this wider view allows the smooth connection of formal and informal logic, as indeed in the informal setting information in many cases appears to be formally inconsistent.

This paper does not aim to describe in detail the technical reformulation of formal logic in argumentation. This can be found in the development of argumentation-based logic in the articles (Kakas et al. 2014; Kakas et al. 2018) which give a precise technical reformulation of classical deductive reasoning in propositional logic (PL) in terms of argumentation. Rather, we will adopt an informal presentation of these technical results, concentrating in exposing the essential elements of this reformulation, as the salient and universal elements of logic through argumentation, across its whole spectrum from formal to informal reasoning. We will therefore avoid the heavy use of

formal terminology and notation, confining ourselves to some parts of the standard language of propositional and first order predicate logic. As mentioned above, readers who are interested in the technical and formal details can consult the articles (Kakas et al. 2014; Kakas et al. 2018).

The rest of the paper is organized as follows. Section 2 presents the basic notions for constructing logical arguments concentrating in the details on the case of formal propositional logic (PL). Then Section 3 describes how using standard semantical notions from formal argumentation theory in AI we can capture strict logical reasoning as a form of reasoning via argumentation. Section 4 discusses the way that the argumentation reformulation of formal logics given in the previous sections extends back into the realm of informal reasoning thus closing the circle of linking the concepts of informal and formal logic. Finally, Section 5 summarizes and discusses the wider significance of the results together with the way in which this opens up new opportunities for the future study of informal logic.

2. Logical arguments

Argumentation is concerned with constructing **good quality arguments** supporting a claim, e.g., supporting a logical conclusion in the case of argumentation-based reasoning. There are therefore two central issues to consider in argumentation: (i) the **construction** of arguments and (ii) the **evaluation** of these as good quality ones. These two aspects are tightly interrelated since in the dialectic process of argumentation, the evaluation of an argument is relative to what other arguments are constructed attacking or defending the initially constructed argument that supports the position of interest.

Let us first consider how arguments would be constructed in an argumentation-based reformulation of formal logical reasoning. As mentioned above, we will follow the approach of argumentation-based logic (Kakas et al. 2014; Kakas et al. 2018), adopting the proof theory formulation of formal logic and building on the connection of proof theory and argumentation. The construction of arguments in informal logic typically follows some accepted **argument schemes** (Pollock 1995; Walton 1995) that would link premises to a conclusion, often called the claim or position, supported by the argument.

Arguments are formed as instantiations of appropriate argument schemes for the particular domain or type of discourse of argumentation, e.g., in the domain of legal argumentation (Prakken 2005) schemes will follow legally accepted patterns of presenting an argument. In other cases of argumentation such as "every day" argumentation we can have more general argument schemes, e.g., a "practical reasoning" scheme for supporting a specific course of action based on the premise that it is used frequently or a "temporal reasoning" scheme for supporting that some property holds based on the premise that it held before at a close previous time.

For formal logical reasoning, which is independent of the content of premises from which we draw conclusions, the appropriate argument schemes depend only on the logical form of the premises. As mentioned above these can be selected from the proof rules of formal logic such as, for example, the well-known *modus ponens* rule. Hence *modus ponens* is seen as a general argument scheme for constructing **logical arguments**, i.e., arguments whose link between premises and supported position can indeed be validated through a precise logical proof in some formal logical system such as classical logic.[1]

This then provides us with the first link in informalizing formal logic. One starts by considering the set of proof rules in a proof system of formal logic as argument schemes. **Arguments** can then be identified with sets of formulae, taking the role of the argument premises, that under some of the proof-rule argument schemes derive a concluding formula, which takes the role of the claim or position of the argument. A natural place from which to select the argument schemes for logical arguments of classical logic is from its proof system of natural deduction (Gentzen 1935). In this system, proof rules have a direct correspondence to argument schemes. For example, the *modus ponens* proof rule in natural deduction, also called implication elimination rule and denoted by

$$\textit{modus ponens:} \quad \frac{\varphi, \varphi \to \psi}{\psi}$$

---

[1] As mentioned above in this paper we will confine ourselves mainly to the case of classical logic and more specifically to classical propositional logic although the ideas presented would apply more generally to other formal logics.

corresponds to an argument scheme that links the set of premises of the form $\{\varphi, \varphi \rightarrow \psi\}$ to the claim of ψ. In other words, we can construct an argument to support the claim ψ based on the premises φ and $\varphi \rightarrow \psi$.

There are two important conditions that need to be applied to this choice of proof-rule argument schemes. The first is that these argument schemes of proof rules need to be considered as **strict schemes**. This means that arguments constructed via these proof rules cannot be defeated by questioning the validity of the chosen proof rules and that such arguments necessarily need to be considered in the argumentation framework that we are constructing. We will see at the end of the paper that once we allow ourselves to reason with contradictory (or inconsistent) premises then this strictness condition is relaxed, and these proof rules can behave as defeasible argument schemes.

The other condition has a more technical nature and requires that the proof rules of *reductio ad absurdum* (RA), or proof by contradiction, are excluded from this initial choice of core argument schemes.[2] In other words, although in classical logic, e.g., in proposition logic, we consider the RA rules on an equal footing with the other proof rules like that of *modus ponens,* this is not so in its argumentation reformulation. To understand this we need to notice that an RA rule contains an element of evaluation of arguments and hence cannot be considered as a primary scheme of construction of arguments. For example, in natural deduction we have the RA rule,

$$\textit{reductio ad absurdum: } \frac{[\varphi \dots \perp]}{\neg \varphi}$$

---

[2] In the proof theory of classical logic we can consider two proof rules of RA, one referred to as *intuitionistic RA*, where the rule leads to the negation, $\neg \varphi$, of the posited hypothesis φ (as shown) and the other, referred to as *classical RA*, where the rule leads to the positive part φ of a negative posited hypothesis of the form $\neg \varphi$. Although using both may be redundant in some cases, the argumentation formulation requires that both of these are excluded as direct argument schemes.

where ⊥ denotes falsity (identified by the conjunction of any formula and its negation), derived here from the hypothetical premise of φ. Given this derivation the rule allows us to conclude $\neg\varphi$.

An RA (or proof by contradiction) rule is made up of two parts: it first shows that a formula posited as a hypothesis is inconsistent when added to the given theory of premises and then uses this to derive the negation of the posited hypothesis. Hence it first evaluates the argument, corresponding to the posited hypothesis, as invalid and then uses this meta-information to support an argument for the negation of the hypothesis. We thus see that the RA rules rely on the semantic quality of the formulae, operating at a meta-level, in contrast to the other proof rules which operate through a *direct support* of their conclusion.

In fact, the main technical task in the reformulation of formal logic in argumentation terms is to recover at the semantic level of argumentation the RA proofs of formal logic. This connection is accomplished by recognizing that the first step in an RA proof, showing that a posited hypothesis is inconsistent, corresponds to showing that any argument based on this hypothesis as a premise is invalid. We will see that arguments involving such formulae in their premise will be recognized as fallacious arguments of a specific type.

Given the above considerations the first part of the construction of logical arguments can be summarized as follows:

- Arguments, **A**, are identified with sets of (propositional) formulae. When we have a given set of premises, i.e., a theory T of given formulae from which we want to draw logical conclusions, we can separate the formulae composing an argument into two parts, **A**=$< T_a, H_a >$, where the first part, T_a is a subset of T and H_a is any subset of formulae, called the **hypotheses** part of A.[3]

---

[3] The use of hypotheses can be viewed as admitting in the argumentation framework a general argument scheme called the **hypothesis argument scheme** used to construct arguments of the form **A**=$< \emptyset, H >$ supporting any (set of) formula(e), H, by simply assuming this (these), i.e., without any explicit support from given premises.

- A subset of proof rules of formal (propositional) logic is selected as strict argument schemes for constructing arguments that support a conclusion. The support is given via a proof of the conclusion from the formulae/premises of the argument under the selected proof rules.
- This subset of chosen proof-rule argument schemes cannot contain an RA (proof by contradiction) proof rule. Technically, in order to fully recover formal (propositional) logic we need to select all proof rules of a complete system for propositional logic except its RA rules.
- The chosen proof rule argument schemes are called **direct argument schemes**. The support of a conclusion, φ, by an argument, $\mathbf{A}=<T_a, H_a>$, is given through a **direct derivation** of φ from $\mathbf{A}$ using only the chosen proof rules. This will be written in the form $\mathbf{A} \Vdash \varphi$ or as $T_a \cup H_a \Vdash \varphi$, where $\Vdash$ denotes derivation under the chosen set of proof rules.

Let us illustrate these ideas with a simple example. Suppose we are given or have learned the following premises about a particular subclass of triangles of interest:

- "any triangle with one angle greater than 100 degrees is not right-angled";
- "any triangle with two angles whose sum is less than 85 degrees is not right-angled".

These can be represented in propositional logic by a theory, T, containing the following sentences (for each triangle in the subclass of triangles that we are considering):

$$one\_angle\_over\_100 \rightarrow \neg right\_angled \quad (1)$$

$$two\_angles\_less\_85 \rightarrow \neg right\_angled \quad (2)$$

Given additional information about any particular triangle we can construct arguments for and against this being a right-angled triangle. For example, if we are given the following additional premise

in T for a particular triangle, which we know has an angle over 100 degrees,

$$\text{one\_angle\_over\_100} \qquad (3)$$

then we can easily construct an argument, $\mathbf{A_1}=< T_1, \{\} >$, with premises, sentences (1) and (3), supporting the conclusion that the triangle is not right-angled, as there is a direct derivation (using the proof rule of *modus ponens*) of this conclusion from the premises in $\mathbf{A_1}$.

Note that for the given theory, T, of premises the only arguments that we can construct to (directly) support that a triangle is right-angled must contain some formula outside T, i.e., what we have called above a hypothesis formula. For example, we can have the argument, $\mathbf{A_2}=< \{\}, H_2 >$, where H_2 consists simply of the formula, right_angled, for the particular triangle at hand, i.e., we claim that the triangle is a right-angled one without any explicit supporting information for this claim.

We will see below the significance of this difference in the type of premises used when we consider the dialectic evaluation process of arguments. As expected, arguments whose premises are entirely drawn from the given theory will be stronger than or preferred over those which are not, allowing for example $\mathbf{A_1}$ to win over $\mathbf{A_2}$ and as a result the given theory, *T*, above to logically conclude that the particular triangle with the property given by (3) is not a right-angled triangle. We will see that recognizing the need for a relative strength between argument schemes and arguments constructed from them plays an important part in connection of formal and informal logic.

3. Logical reasoning as dialectic argumentation

Constructed arguments are judged against other arguments and in particular against their counter-arguments, such as arguments supporting positions that are in conflict with them. In general, an argument is deemed to be of good quality if it can defend against its counter-arguments, e.g., by attacking its counter-arguments in return. This evaluation of an argument is a **dialectical process** of considering counter-arguments and then arguments defending against these counter-arguments. Do which counter-arguments to take into account and what the allowed forms of defense are vary according

to the application domain of argumentation we are considering? For example, in the case of strict formal logical reasoning, an argument must defend against all possible counter-arguments whereas in informal logical reasoning only a (flexibly small) subset of counter-arguments may be considered.

In the dialectic process of argumentation, given a position of interest we can distinguish **pro** and **con arguments**, i.e., arguments that support the position and arguments that oppose the position. This relative characterization of the arguments is typically based on some form of conflict or incompatibility in the domain of discourse. In general, when we are not interested in taking sides, we consider a symmetric relation between such arguments, calling them **counter-arguments with respect** to each other or arguments **attacking** each other. In artificial intelligence the attacking relation in an argumentation framework often contains more information than simply this symmetric incompatibility of the arguments involved (Bondarenko et al. 1993; Dung 1995; Modgil and Prakken 2013). This extra information, as we will see below, pertains to the relative strength or preference of the arguments involved in the attack.

In formal logic, incompatibility or contradiction is captured via the conflict between a formula and its negation. Hence if one argument supports, i.e., directly derives, a formula, φ, while another supports its negation, $\neg\varphi$, these two arguments will attack each other. More precisely, for formal propositional logic this (symmetric) attack relation can be captured through the joint **direct derivation** (denoted here by $\Vdash$) of an inconsistency, $\bot$, i.e., the derivation of falsity using only the direct proof-rule argument schemes that we have chosen. Hence any two arguments, $\mathbf{A}_1 = <T_1, H_1>$ and $\mathbf{A}_2 = <T_2, H_2>$, **attack each other** if and only if we have that $\mathbf{A}1 \cup \mathbf{A}2 \Vdash \bot$, i.e., $(T_1 \cup H_1) \cup (T_2 \cup H_2) \Vdash \bot$.

To complete the reformulation of formal logical reasoning as a dialectic argumentation process we also need to define the way in which arguments can defend against other arguments. For the moment, let us simply consider the defense relation between arguments to be the same as (or a subset of) the attack relation and return to this after we have defined the central notion of validity or acceptability of arguments.

3.1 Validity or acceptability of arguments

Arguments are **acceptable** when they exhibit a good dialectic quality, namely that they can **"defend against all attacking arguments"**. Analogously, an argument is **non-acceptable** if there is at least one **counter-argument** that it **cannot be defended against**.

The important observation here is that an argument can come to its own defense to ensure its acceptability or that an argument can contribute to its own demise rendering itself non-acceptable. In fact, by defending against its counter-arguments an argument typically takes a position that contributes in rendering its counter-arguments non-acceptable. Similarly, a non-acceptable argument, **A**, can be so because it helps to render one of its counter-arguments, which is in general non-acceptable, into an acceptable argument in the context of **A**, i.e. when we are considering accepting **A**. Such arguments, **A**, are **self-defeating** or **self-refuting** arguments (see, e.g., Pollock 1991 for a related notion of self-defeating arguments). We will see that these arguments play an important role in our reformulation of formal logic in terms of argumentation, as they correspond to indirect logical proofs by *reductio ad absurdum*: they encapsulate at a more general level of validity of argument the notion of proof by contradiction in formal logic.

To turn this specification of acceptability and non-acceptability of arguments into a precise definition we need to additionally notice that the defense against any counter-argument must also be required to be acceptable and, importantly, acceptable within the context of the original argument that we want to be acceptable. Thus the notion of **acceptability** of arguments is a *relative notion*, A_1 **is acceptable w.r.t.** A_0, for any two arguments, A_0 and A_1. This is recursively specified by:

> A_1 **is acceptable w.r.t.** A_0 if for any argument, **B**, attacking A_1 there exists a defending argument, A_2, against **B**, such that A_2 **is acceptable w.r.t.** (A_0 extended[4] by A_1).

---

[4] This extension forms coalitions of arguments. For simplicity, we can consider extension to be simply given by set union.

The base case of this recursive definition of acceptability is given by having any argument A_1 contained in A_0 be acceptable w.r.t. A_0, reflecting that A_0 and the formulae that it contains are considered as an initially acceptable context of arguments. An argument A is then said to be **dialectically acceptable** or simply **acceptable** if and only if it is acceptable relative to the empty set, i.e., **A is acceptable w.r.t. {}** holds.

For the non-acceptability of arguments, i.e., arguments that cannot defend against all their attacking arguments, we can analogously, define a relative notion of non-acceptability of arguments recursively that captures this as follows:

> A_1 **is non-acceptable w.r.t.** A_0, if there exists an attacking argument **B** against A_1 for which any possible defense, A_2, against **B** is **non-acceptable w.r.t. (A_0 extended by A_1)**.

The base case of this is when we have an attack **B** that is contained in A_1, i.e., when A_1 is explicitly non-acceptable because it attacks itself: it is explicitly self-defeating. For example, inconsistent formulae are always non-acceptable because they are self-attacking.
Given these relative notions of acceptability and non-acceptability, an argument **A** is said to be **dialectically acceptable (resp. non-acceptable)** or simply **acceptable (resp. non-acceptable)** if and only if it is acceptable (resp. non-acceptable) relative to the empty set, i.e. **A is acceptable (resp. non-acceptable) w.r.t. {}** holds.

Let us now return to the issue of the specification of the defense relation. Informally, to capture this we will consider a notion of relative strength or preference between logical arguments allowing stronger arguments to defend against weaker ones but not vice-versa and so stronger arguments will "win over" the weaker arguments. The **defense relation** between arguments normally captures the relative strength or preference between arguments. In general, an argument **defends against** a counter-argument if it is preferred over the attacking argument or they are non-comparable in preference. In other words, the defending argument is not of lower preference than the argument it defends against.

The preference and its defense relation in many domains of argumentation comes from domain specific information. For example, in legal argumentation this could reflect legal precedence such as *lex-*

posterior where later laws are generally stronger than earlier laws. For the general framework of argumentation-based logical reasoning as applied for the case of propositional logic (Kakas et al. 2014; Kakas et al. 2018) the preference and ensuing defense relation is minimal and consists of two elements:

– Arguments which are entirely made out of premises in the given theory *T* are **strictly preferred** over arguments that contain hypothetical sentences and thus can be defended against only by other arguments that also consist entirely from premises in *T*. Arguments that contain hypotheses cannot defend against such arguments made entirely from the given theory *T*.

– A formula φ and its complement, φ^c, are **equally preferred**. We are free to choose equally between the two (provided that one is not a direct consequence of the given theory *T*), with this choice allowing us to take the side appropriately needed to defend against attacks. Hence any argument **A** containing a hypothetical formula φ can be defended against by the argument **D**=$<\{\},\{\varphi^c\}>$ (also denoted by **D**=$\{\varphi^c\}$), i.e., **A** is defended against by **undermining** one of the (hypothetical) premises on which it rests.

Note that the second element of defense does not mean that for any formula we must choose either that formula or its negation—in the same fashion in which we build two valued models in logic—since it could be that for some particular formula both options are non-acceptable arguments, as we find for example in the case of some logical paradoxes. In such cases the premises, *T*, will typically be (classically) inconsistent but as we will see in the next section this does not have a trivializing effect on the argumentation formulation of logic and indeed argumentation logic then deviates from formal classical logic.

When the given premises, *T*, are consistent the first element of defense means that attacking arguments that are made entirely from *T* cannot be defended against. Hence an argument that is attacked by an argument made entirely of premises in *T* cannot be acceptable. Similarly, an argument **A** made entirely from *T* is attacked only by arguments containing hypothetical formulae and so these can always

be defended against by **A** itself. Therefore, such arguments drawn entirely from *T* are always acceptable.

3.2 Example continued

To illustrate these semantic notions of acceptability and non-acceptability let us return to our triangle example given above. We can easily see for any particular triangle which has one angle greater than 100 degrees (i.e., *one_angle_over_*100 is included in the premises, *T*), the argument **A**$_1$=<{1, 3}, {}> is acceptable, since it is entirely drawn from the premises in our given theory, and hence the conclusion that the triangle is not right-angled is **acceptably supported**. In fact, the opposite conclusion cannot be supported by any acceptable argument. This is because any argument whose direct conclusion is *right_angled* will be attacked by **A**$_1$ (a subset of the given premises) and since the premises are consistent this cannot be defended against.

When we have a situation like this where a formula is acceptably supported, and its complement is not acceptably supported, we will say that this formula is a logical conclusion under argumentation of the premises *T*. Hence ¬*right_angled* is a logical conclusion under argumentation as we would need if argumentation logic is to be equivalent to formal propositional logic.

For an illustration of a more complex case of the dialectic argumentation process and how this captures the formal logical conclusions of propositional logic, let us consider an extension of the above triangle example where we have learned that for the particular subclass of triangles that we are considering the following additional statement holds:

– "if a triangle does not have two angles whose sum is less than 85 then it has one angle over 100 degrees".
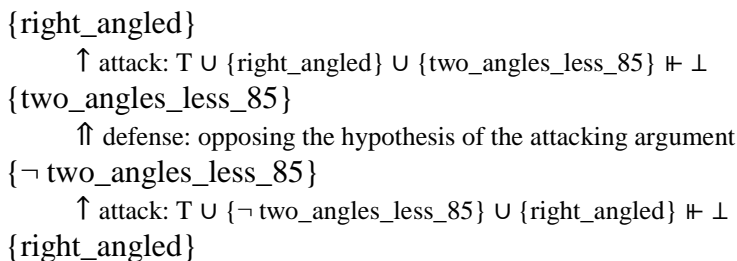
This is represented formally as:

$$¬two\_angles\_less\_85 → one\_angle\_over\_100 \quad (4)$$

Hence we are now considering the theory, *T*, consisting of sentences (1), (2), and (4). Let us see how in the framework of

argumentation-based logical reasoning, *right_angled* is judged as non-acceptable. We will see below that this would then (skeptically) entail in argumentation that ¬*right_angled* holds, as indeed is the case for classical propositional logic. The position of *right_angled* can only be supported by arguments that contain this as a hypothesis or directly derive this from hypotheses. Then the non-acceptability of such arguments can be determined by considering the counter-argument consisting of the premise (2) from *T* and the hypothesis {*two_angles_less_*85}. The dialectical process of argumentation that shows that this attack cannot be defended against is depicted in Figure 1 where for simplicity we only show the hypothesis part of the arguments involved.[5]

{right_angled}
      ↑ attack: $T \cup \{right\_angled\} \cup \{two\_angles\_less\_85\} \Vdash \bot$
{two_angles_less_85}
      ⇑ defense: opposing the hypothesis of the attacking argument
{¬ two_angles_less_85}
      ↑ attack: $T \cup \{\neg two\_angles\_less\_85\} \cup \{right\_angled\} \Vdash \bot$
{right_angled}

Fig. 1. Dialectic process of argumentation for determining the non-acceptability of *right_angled*, given $T = \{(1), (2), (4)\}$, in order to determine the classical entailment of ¬*right_angled* from *T*. Arguments are shown only by their hypotheses as indicated in brackets.

The informal reading of this figure is as follows: The argument of supposing that a triangle (in this subclass) is right-angled is attacked by the hypothesis that the triangle might have two angles whose sum is less than 85 degrees. The canonical objection or defense to this counter-argument is to assume instead that the triangle does not have two angles whose sum is less than 85 degrees. But this defense is in conflict with the original argument it is meant to help defend, as together they directly derive through (2) and (4) an inconsistency. The only possible defense is attacked by the original hypothesis of *right_angled* that it is trying to defend. Hence *right_angled* is self-

---

[5] For simplicity we also assume here that $\Vdash$ contains only the *modus ponens* proof-rule argument schema.

defeating and thus non-acceptable. In dialectic terms we can read this as a debate between pro and con arguments as follows:

> **Pro:** I believe that the triangle is right-angled.
> **Con:** Not so, as the triangle can have two angles whose sum is less than 85 degrees.
> **Pro:** That is only a supposition and I can assume that this is not so, i.e. that it does not have two angles whose sum is less than 85 degrees.
> **Con:** But if the triangle does not have two angles whose sum is less than 85 degrees then it will have one angle over 100 and so it cannot be right-angled.

The con side "wins the argument" as the only possible defense against the objection that it raises, namely that the triangle does not have two angles whose sum is less than 85 degrees, cannot be accepted in the context of the original argument that we are trying to accept. In other words, the argument of *right_angled* turns its needed defense, ¬*two_angles_less_85*, invalid or non-acceptable. This defense is objected to by the very argument that it is trying to defend. Hence although in general there is a defense against the objection (by taking the opposite view) this is not possible in the context of the particular argument that we wish to be acceptable. This shows the importance of considering the notion of validity or acceptability of arguments as a relative notion, in the context of arguments that we are taking for granted, either in general or in the specific case of acceptability (or non-acceptability) that we trying to decide upon.

3.3 Rationality and logical entailment

Defenses against attacks to an argument must "hold well together" in the sense that they need to be conflict free or in other words they should not contain an internal attack or counter-argument relation between them. This is the basic **rationality property** of the dialectic argumentation semantics. It can be stated in simple terms as follows: "none of the defenses should form a counter-argument against any of the other defenses and of course against the original argument of interest". In the example above the defense proposed in the third step is a counter-argument against the originally proposed argument and

hence the rationality property is violated. In general, for any acceptable argument there must exist a set of defenses against all its attacks that satisfies the rationality property of being conflict free.
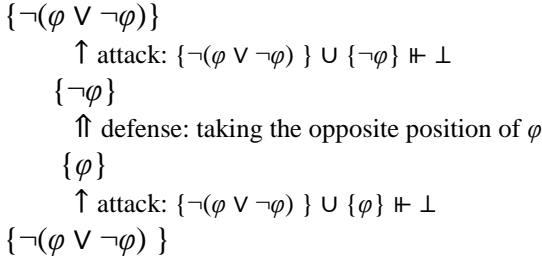
This rationality property of the set of defenses points to the connection between acceptability of arguments and the formal logical notion of **satisfiability** of the formulae composing the argument. In fact, in the case where the given premises, *T,* are consistent and we have a classical model for the set of formulae in an argument, *A*, then, it is easy to see that we can choose the defenses for *A* from the set of formulae that are made true in this model and hence *A* would be acceptable. In other words, satisfiability implies acceptability and vice versa. This means that when the (propositional) premises, *T*, are classically consistent then formal logical entailment, in the classical sense of truth in all models, coincides with the notion of **skeptical conclusion** in argumentation defined as follows:

> A formula *φ* is a **skeptical conclusion** of argumentation if and only if *φ* is dialectically acceptable and ¬*φ* is dialectically non-acceptable.

The non-acceptability of the negation of the concluding formula means that this is not possible under any circumstances and therefore the positive conclusion is absolute. Hence it is called **skeptical** in the sense that it is a **very strong** conclusion that even skeptics would draw.

To illustrate further the equivalence between formal logic and argumentation logic let us consider the example of the logical entailment of the **excluded middle law**, that for any formula *φ* the formula, *φ* ∨ ¬*φ,* is a tautology, i.e. it is a logical conclusion from the empty set of premises (and consequently from any set of premises). In a similar way to the triangle example above we can argue that the negation of the excluded middle law is non-acceptable. As we do not have any premises, the only way to support the negation of the law is through some set of hypothetical formulae which directly derive this. Figure 2 shows the way in which such arguments are not acceptable by considering the prototypical such argument, $<\{\}, \{\neg(\varphi \vee \neg\varphi)\}>$, i.e., where we argue for it by explicitly hypothesizing it. This is attacked by ¬*φ* since together the two formulae lead to falsity, under the "disjunction introduction rule" direct proof

rule (argument scheme) whereby a formula φ can form the premise for an argument supporting the disjunction conclusion $\varphi \vee \psi$ for any formula ψ.

$\{\neg(\varphi \vee \neg\varphi)\}$
    $\uparrow$ attack: $\{\neg(\varphi \vee \neg\varphi)\} \cup \{\neg\varphi\} \Vdash \bot$
  $\{\neg\varphi\}$
    $\Uparrow$ defense: taking the opposite position of φ
   $\{\varphi\}$
    $\uparrow$ attack: $\{\neg(\varphi \vee \neg\varphi)\} \cup \{\varphi\} \Vdash \bot$
$\{\neg(\varphi \vee \neg\varphi)\}$

Fig. 2. Dialectic argumentation for the excluded middle law.

This attack of $\neg\varphi$ can only be possibly defended by φ but this is attacked by the argument of $\neg(\varphi \vee \neg\varphi)$ (again via the disjunction introduction rule) and hence $\neg(\varphi \vee \neg\varphi)$ is not acceptable. This together with the fact that $\varphi \vee \neg\varphi$ is acceptable implies that $\varphi \vee \neg\varphi$ is skeptically entailed.

Technically, the correspondence between the logical systems of classical logic and argumentation logic rests on a mapping of the dialectic argumentation process (such as the examples we have seen in Figures 1 and 2) to formal logical proofs in the proof system of natural deduction (Gentzen, 1935). This mapping is onto a restricted form of natural deduction derivations where the use of *reductio ad absurdum* is limited in a particular way. This restriction corresponds to the fact that self-attacking arguments cannot affect the acceptability of arguments that they attack, and so when we apply a proof-by-contradiction rule in natural deduction, reaching a contradiction must necessarily involve the hypothesis posited for the proof to be admitted. In this way, argumentation re- covers semantically the effects of proof by contradiction in logical entailment and gives us a formal derivation for classical logic that is logically sound and complete with respect to the semantics of classical logic, when the given theory, *T*, is classically consistent. The interested reader can see, in (Kakas et al. 2018), the technical details of this correspondence with natural deduction derivations and how this gives a way to understand

the precise equivalence of argumentation logic to formal classical propositional logic.

3.4 Formal syllogistic reasoning

Aristotle introduced his systematic study of logic through syllogistic reasoning. In modern form syllogisms can be seen as canonical forms of classical reasoning within a subset of first order predicate logic. In this subsection we will present the way in which the argumentation-based reformulation of formal logic applies to syllogistic reasoning. We will briefly describe how syllogistic reasoning is captured through argumentation in the same "strict limit" described above for propositional reasoning. We will illustrate this with a few representative examples chosen out of the 64 possible syllogisms introduced by Aristotle.

So let us start with the *Barbara*[6] syllogism where from the statements that "All As are Bs" and "All Bs are Cs" we can conclude that "All As are Cs". The premises of this syllogism are naturally represented in first order predicate logic by the following statements, T:

$$\forall X(A(X) \rightarrow B(X)) \quad (5)$$
$$\forall X(B(X) \rightarrow C(X)) \quad (6)$$

To check that the syllogism holds, i.e. entailed via argumentation, we need to consider the acceptability of arguments supporting $\forall X(A(X) \rightarrow C(X))$ and $\neg\forall X(A(X) \rightarrow C(X))$ and show that the first formula is acceptable whereas the second one is not acceptable.

The formula $\forall X(A(X) \rightarrow C(X))$ is a *direct consequence* from the premises, T, i.e., it is supported by a set of direct argument schemes corresponding to natural deduction proof rules for full first order classical predicate logic not containing the proof rule for *reductio ad absurdum*.[7] As we have seen above, formulae which are direct

---

[6] We are using here the standard nematology given to the syllogisms by the medieval scholars.

[7] Informally, to see how $\forall X(A(X) \rightarrow C(X))$ is a direct consequence we consider an arbitrary object named by a term, t; start from the hypothesis, $A(t)$; and with a double application of the *modus ponens* argument scheme (with the two premises in T) we can form an argument that supports $C(t)$. Then the direct argument

consequences are acceptable and their negation is not acceptable and hence, as required, $\forall X(A(X) \rightarrow C(X))$ is skeptically entailed.

Another example of an Aristotelian syllogism (called *Cesare*) has as given premises the two statements: "All As are Bs" and "No C is B", from which we can conclude that "No A is C". These premises are again naturally represented in classical first order logic by the statements, *T*, given by:[8]

$$\forall X(A(X) \rightarrow B(X)) \quad (7)$$
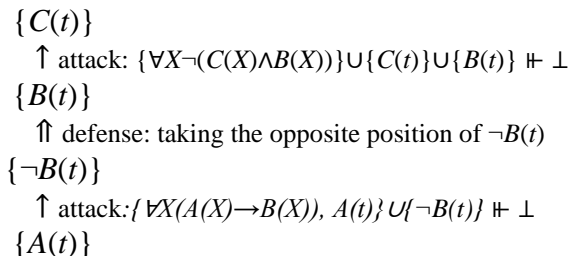$$\forall X \neg (C(X) \wedge B(X)) \quad (8)$$

In this example the syllogistic conclusion, $\forall X \neg (A(X) \wedge C(X))$, is not a direct consequence under predicate logic. At least one application of *reductio ad absurdum* is required. One way to see this is to assume that $A(t)$ holds for some arbitrary term *t* and examine whether $C(t)$ can hold or not. To decide on $C(t)$, we do not have a direct proof for it or its negation, and indeed we need to apply *reductio ad absurdum*; i.e., we assume $C(t)$ holds as a given premise and note that this would lead to a contradiction. This happens because from the first premise, (7), and $A(t)$, we also have $B(t)$ which through the second premise (8) contradicts $C(t)$. Hence through *reductio ad absurdum* $\neg C(t)$ holds, as required by the desired syllogistic conclusion.

In terms of argumentation, this proof by contradiction is captured by realizing that $A(t)$ renders $C(t)$ non-acceptable: $C(t)$ is non-acceptable w.r.t. $A(t)$. To see this let us assume that $A(t)$ is part of the given premises *T* (so that we are in the context of $A(t)$) and consider the acceptability of $C(t)$ and how this might be related to the classical proof by contradiction described above. We have the following dialectic argumentation tree, shown in Figure 3.

---

scheme corresponding to the proof rule for the introduction of implication gives support for $A(t) \rightarrow C(t)$. Finally, the argument scheme for the introduction of the universal quantifier gives support for the general formula $\forall X(A(X) \rightarrow C(X))$ since *t* is a completely arbitrary term.

[8] Another choice to represent the second premise is $\forall X(C(X) \rightarrow \neg B(X))$. This does not change our analysis of the argumentative reasoning that captures the syllogistic conclusion.

$$\{C(t)\}$$
$\quad \uparrow$ attack: $\{\forall X \neg (C(X) \wedge B(X))\} \cup \{C(t)\} \cup \{B(t)\} \Vdash \perp$
$$\{B(t)\}$$
$\quad \Uparrow$ defense: taking the opposite position of $\neg B(t)$
$$\{\neg B(t)\}$$
$\quad \uparrow$ attack: $\{\forall X(A(X) \rightarrow B(X)), A(t)\} \cup \{\neg B(t)\} \Vdash \perp$
$$\{A(t)\}$$

Fig. 3. Argumentative reasoning in the *Cesare* syllogistic example.

The important observation is that $A(t)$ renders a defense that $C(t)$ necessarily needs, namely $\neg B(t)$, non-acceptable. Hence $C(t)$ is non-acceptable in the context of $A(t)$ and thus it is not possible to have together $A(t)$ and $C(t)$ for any arbitrary term t. From this we can then draw the general syllogistic conclusion $\forall X \neg (A(X) \wedge C(X))$, i.e., "No A is C".

Similarly, we can analyze and capture all (24) formally valid syllogisms in terms of argumentation. We will see in the next section that by extending the argument schemes and the preference relation between them we can also capture within the same framework human syllogistic reasoning, i.e., conclusions drawn by humans irrespective of whether these are formally valid or not under classical logic. In other words, we will see how we can uniformly capture formal and informal syllogistic reasoning within the single framework of dialectic argumentation.

4. Beyond classical reasoning

Classical propositional and first order predicate logics are not equipped or designed to deal with contradictory information. When the given premises, *T*, are inconsistent, formal classical reasoning collapses in the sense that every formula is trivially entailed. In contrast, argumentation is concerned exactly with how to deal with contradictory information and conflicting positions.[9] Hence the

---

[9] Aristotle in his first books of the *Organon* studies dialectic argument and valid schemes of argumentation before introducing the strict form of syllogistic reasoning.

argumentation-based reformulation of formal classical reasoning that we have described above for the case of consistent premises, *T*, would or should carry through when *T* becomes **inconsistent**.

Consider for example that we find out that some triangle in our subclass of triangles is in fact a right-angled triangle. Adding this fact to the premises {(1), (2), (4)} that we already have turns the set of premises classically inconsistent. Nevertheless, the argumentation-based formulation of logic that we have described above will not trivialize under these inconsistent premises. For example, it would skeptically conclude that this particular triangle is right-angled, as we might expect since we have "explicit" information for this, without also concluding that it is also not right-angled, as propositional logic does.

The dialectic argumentation proof in Figure 1, that gave us earlier the conclusion that the triangle is not right-angled, now changes as we have another way to defend against the attack(s) containing the hypothesis *two_angles_less_85*. We can now defend using the premise *right_angled* that, as we have explained above, is preferred over arguments that contain hypotheses as it is made purely of sentences in the given premises T. On the other hand, any argument supporting ¬*right_angled* will be attacked by the argument made purely from the premise *right_angled*. This cannot be defended against since there is no direct or explicit information to its contrary in the premises. Hence *right_angled* is acceptable and any argument supporting ¬*right_angled* is not acceptable, thus *right_angled* is skeptically concluded.

The extended theory considered above, i.e. {(1), (2), (4), *right_angled*}, although classically inconsistent it is still directly consistent, i.e., there is no conflict that is directly derivable. It is not possible to derive using the direct proof rules alone (i.e., without the use of a *reductio ad absurdum* rule) both some formula and its negation. Hence there are no two arguments made entirely from the given (classically inconsistent) premises that are in direct conflict and hence attack and defend each other. This means that these chosen direct proof rules still behave as strict argument schemes. But when the given premises are directly inconsistent then the direct proof rules will behave as **defeasible argument schemes**: they construct arguments but the formulae that they support do not

necessarily follow from the given premises. An example of this would be a case where we have two explicit opinions (from equally trusted sources) about the same triangle, one stating that it is right-angled and the other stating that it is not right-angled. Then each of these premises would defend against each other and so both arguments would be acceptable and therefore there will be no skeptical conclusion for whether the triangle is or is not right-angled.

The strict conditions on the argumentation framework that we have imposed so that we can match the strict reasoning of formal logic can be further relaxed by allowing the premises themselves to take a defeasible nature, e.g. implications to represent only a "normally or mostly" nature of associating their conditions with their conclusions. This is the case for example of non-monotonic logics, e.g. default logic (Reiter 1980), studied widely in AI. Importantly, in the 1990s it was shown that most of these logics can be reformulated in terms of argumentation, see, e.g., (Bondarenko et al. 1993).

For such defeasible reasoning the defense relation is enriched with relative and often domain dependent preferences amongst this defeasible knowledge by rendering some arguments stronger and hence these arguments are able to defend against (some of) their counter-arguments but not vice-versa. This is particularly appropriate when we consider the informal form of **commonsense reasoning** (Kowalski 2011; Stenning and van Lambalgen 2008a) where people normally reason within a context and where the commonsense knowledge is in this form of loose or default associations. General or individual **human biases** give preference to some of these statements and we can then understand common sense reasoning in argumentation terms in the same way we have expressed formal logical reasoning[10] within argumentation.

Let us discuss this further and illustrate it by considering a particular example of such human reasoning as we find in the task of text

---

[10] We already saw a case of this when, in our triangle example, we considered for the particular triangle for which we had explicit information that this was a right-angled one. In human reasoning there is a natural preference bias of explicit information over other conflicting information not resting on explicit information. This human bias is supported by several empirical observations from studies in the cognitive psychology of reasoning.

comprehension such as when people read and comprehend a story. Consider the following start of a story:

- Mary was very busy at the office.
- Her phone rang.

Did Mary want to answer the phone? An argumentation process to answer this question in favor of not wanting to answer would require constructing an argument supporting this conclusion and acceptably defending it against its counter-arguments. One such argument might use as its premises the commonsense knowledge that "normally, when people are busy they do not want to be disturbed", and "normally, when people do not want to be disturbed they do not want to answer their phone". Grounding these with the information that Mary was busy we get an argument supporting the conclusion that Mary does not want to answer the phone. Let us call this the "busy argument". This argument is attacked by another argument based on the commonsense knowledge that "normally, people want to answer their phones". Let us call this the "general argument". But the author by including the qualification of "very" awakens a preference of the "busy argument" over the "general argument". The "busy argument" can then defend against the counter-argument but not vice versa, and the conclusion of "not wanting to answer the phone" is drawn. The preference and its effect could be reinforced by the author adding an extra sentence in the story, e.g. starting the story instead as follows, where the second sentence now provides an argument explicitly against wanting to answer the phone:

- Mary was very busy at the office.
- She did not want to be distracted.
- Her phone rang.

Of course, this may not be the way all readers would reason. Other factors come into play when we are reading stories. A "lazy reader" would not want to draw conclusions even when pressed to answer questions until the author makes intentions clear. On the other hand an "attentive reader" may use other types of commonsense knowledge, e.g., meta knowledge about the structure of stories, to

build an argument for supporting Mary wanting to answer the phone based on "normally, calls in a story are significant events" and "normally the actor would be expecting the call". Such an "expect argument" may be non-comparable in strength with the "busy argument" and hence both will defend against each other and so both would be acceptable and hence no conclusion would be clearly (skeptically) drawn. Let us assume that the story continues as follows:

– Mary was very busy at the office.
– She did not want to be distracted.
– Her phone rang.
– It was her mother phoning.

Does Mary want to answer the phone? Will Mary answer the phone?

Now a new argument, let us call it the "mother argument", enters the arena, based on the commonsense knowledge that "normally, people want to answer calls from their mothers". In general, the bias giving preference to this argument may not be strong enough to overcome the earlier "busy argument" supporting the contrary position of not wanting to answer the phone. For some individuals it may be that the new argument is stronger and therefore their conclusion, of not wanting to answer, will be revised. Continuing the story with the following sentences:

– Mary's mother fell ill last week.
– She was still (very) ill in the hospital.

Now the general preference bias for the "mother argument", based on the qualified knowledge that Mary's mother is ill, will in general reinforce this argument making it stronger than the "busy argument" and we will have more readers revising their conclusion (or making up their mind) as now only arguments for wanting to answer the call are dialectically acceptable.

Using this dialectic process of argumentation based on background common sense world knowledge it is possible to construct, along the lines described above, what in cognitive psychology is

called the **mental or comprehension model** (Johnson-Laird, 1983) and understand the process of revision of this model as the story unfolds through the non-monotonicity of reasoning through argumentation (Diakidoy et al., 2014, 2015).[11]

It is interesting to note that once we lock ourselves into a specific comprehension model (at any particular point in the story) the conclusions that this model contains can be seen as formal logical conclusions of the explicit information in the story and the world knowledge that we have used to construct the model. In other words, the informal logic dialectic argumentation that we use to construct the comprehension model is one and the same as the formal logic dialectic argumentation process when projected down on the specifics used for constructing the model. Hence, the informal logic of comprehension has a formal logic interpretation under the strict form of argumentation for capturing formal logic, as we have described in the previous sections. Argumentation thus glues together informal and formal logic in both directions.

Finally, we note that in dynamic environments like the comprehension one or indeed a debate environment, where new information comes that we cannot anticipate *a priori*, argumentative decisions (Paglieri 2013), such as decisions on what arguments to construct or what attacking arguments to react to and when to terminate the argumentation process, would need to be considered to allow a fully flexible argumentation framework—an extreme form of informal logic on the other side of the extreme strict formal logic.

4.1 Human syllogistic reasoning

Reasoning beyond or outside the formal boundaries of classical logic can be found abundantly in human reasoning. Just as in the case of comprehension, cognitive psychologists and behavioral economists have observed that human reasoning across many and different situations deviates from classical reasoning (see, e.g., Evans et al. 1993; Stenning and van Lambalgen 2008b; Thaler 1991; Tversky and Kahneman 1974 for surveys of these empirical results).

---

[11] The construction of a comprehension model depends also on other factors, e.g., that of coherence, where only conclusions in the main thread of the story are considered, but these are extra-logical processes outside the scope of this paper.

Even in cases where humans are asked to reason within a formal setting, such as the one of syllogistic reasoning, it has been observed that they are not "classical reasoners". In (Khemlani and Johnson-Laird 2012) the authors have given a comprehensive review of data on human syllogistic reasoning gathered from several experiments since 1978. This and other recent experiments confirm the observation that although humans will on the whole recognize and draw formally valid conclusions there are significant deviations. Humans can miss valid conclusions, but more importantly, humans will systematically draw conclusions that are not formally valid. A simple case of this can be seen in the following syllogism where the premises have the example form:

$$\text{\textit{Some artists are not bakers.}} \quad (9)$$
$$\text{\textit{All bakers are chemists.}} \quad (10)$$

Representing these in the standard way in first order predicate logic as follows:

$$\exists X(artist(X) \wedge \neg baker(X)) \quad (11)$$
$$\forall X(baker(X) \rightarrow chemist(X)) \quad (12)$$

we can see that there are no formally valid conclusions within the syllogistic language, i.e., there is no syllogistic categorical sentence relating the predicates "artists" and "chemists", that is a formally valid conclusion of these premises. However, a significant amount of human participants answered *Some artists are not chemists* (46%) and only 20% concluded *No valid conclusion* (Khemlani and Johnson-Laird 2012). A recent Syllogism Challenge competition, set in 2017 (see https://www.cc.uni-freiburg.de/syllogchallenge), aimed to examine how well cognitive models of human reasoning can account for the observed data of human syllogistic reasoning.

Within the framework of argumentation-based logical reasoning, presented in the previous sections, we can accommodate human syllogistic reasoning by simply extending the set of argument schemes

that we assume humans apply.[12] In other words, together with the formal argument schemes we have used above, drawn from direct proof rules of formal logic, we can now include additional *informal argument schemes*. These are cognitively justified based on observations of human reasoning behavior, independently from the observations of human syllogistic reasoning. For example, we can include an argument scheme based on the cognitive principle of "maxim of information" (Grice 1975) that humans provide and assume maximum information in their premise statements, e.g. statements that they exchange between them. This maxim provides support for the following informal argument scheme within syllogistic reasoning:

$$\frac{\exists X(a(X) \land b(X))}{\exists X(a(X) \land \neg b(X))}$$

I.e., from the premise statement "$\exists X(a(X) \land b(X))$" (corresponding to the syllogistic statement that "Some As are Bs") we can support the position that "$\exists X(a(X) \land \neg b(X))$". Hence, for example, from the given premise that *Some bakers are chemists*, we can also support the position that *Some bakers are not chemists* by applying this new argument scheme. The "maxim of information" states that we make our exchange of information as informative as possible and therefore we can assume that not all bakers are chemists, as otherwise, we would have expressed the premise that *All bakers are chemists* instead of *Some bakers are chemists*. In this way, the above informal argument scheme is cognitively justified by the independently observed "maxim of information" in human behavior.

Adding such cognitively valid informal argument schemes to the formal argument schemes, that are used to reformulate classical logic, it was shown that a hybrid framework of formal and informal reasoning, but unified through the reasoning form of dialectic argumentation, covers well the observed human syllogistic reasoning behavior. In the syllogistic challenge mentioned above, where entries

---

[12] Note that in syllogistic reasoning the given premises are classically consistent. Hence the extension of formal syllogistic reasoning afforded by the argumentation view is on another dimension from the one presented above of extending the reasoning in the realm of inconsistent premise information.

were tested on hitherto unseen empirical data taken from 140 people, such an argumentation-based approach matched well that data and performed well with respect to the other entries in the challenge.

Note that once we have a hybrid argumentation-based logical framework that includes new informal argument schemes the question arises of how strong are these new schemes? A partial answer to this, given implicitly by the results of the challenge, is simply that the same priority preference as in the case for the reformulation of formal logic in terms of argumentation applies, namely that these argument schemes are stronger than the "hypotheses scheme", i.e. the scheme where we support a conclusion without any premises on which to rest this claim. In addition, any argument scheme that directly or explicitly provides or supports factual information is stronger than arguments which do not do so. Again this new preference is cognitively valid as humans are observed to have greater trust to factual information rather than information derived through some chain of construction. It is important to note that there is no relative preference between formal and informal schemes. In this sense, within such hybrid reasoning frameworks, the formal argument schemes lose their strictness, they become defeasible like the informal ones!

5. Discussion and conclusion

We have described a smooth connection between formal, logically strict, reasoning and informal human reasoning within the same framework of dialectic argumentation. Formal classical logic with its strict logical deductive reasoning can be given an argumentative formulation in the same argumentative terms as that in which informal logical reasoning is normally described. "Informalizing formal logic" in this way provides a way to unify these two worlds of formal and informal logic, normally considered as very different, under the same conceptual framework of argumentation.

Reformulating classical logic and other formal logical systems in terms of argumentation has been investigated in several previous works, notably that of dialogical logic (Lorenz 2001; Ruckert 2001). Nevertheless, the emphasis of these works is to remain within the formal realm of logic, and they are not primarily concerned with the

wider task of unifying formal and informal logic. In contrast, our approach is motivated primarily from work on argumentation that comes from studies of informal human reasoning in artificial intelligence. The aim is to form a two-way link between formal and informal logic and we have seen that by varying the degree of flexibility within a basic argumentation framework drawn from AI and its dialectic process, we can move from formal logic to informal logic and back.

Let us summarize here the essential elements of argumentation that help give a unified framework of formal and informal logic.

- The unification is based on the "standard" conceptual framework of dialectic argumentation with the notions of attacking and defending arguments. The defending relation between arguments is regulated by a relative preference or strength between the argument schemes and their application.
- In formal logic this preference relation between arguments has a relatively simple form, namely that arguments based on given premises are stronger than arguments that are hypothesized, but this plays an essential role in the argumentation reformulation. Importantly, it provides the link to informal logic where preferences can capture the various human biases that are prevalent in their reasoning.
- In our reformulation of formal logic within argumentation the two opposing views of any formula as a conclusion and its negation are treated separately where the conclusion needs to be argued as acceptable and its negation needs to be *separately argued* as non-acceptable. Although, within the realm of consistent premises it is possible to avoid this separation of concern, it becomes important when we have inconsistent premises, as it is in many cases of informal reasoning where information from the environment can be contradictory.
- Another significant element of the argumentation approach that we are using is the relative nature of the acceptability or validity of arguments. This allows us to "reason in context" and to thus recognize quality or not in arguments in different situations, as we find typically in informal reasoning.

– On a more technical level, the relative nature of acceptability and in particular of non-acceptability is linked to self-defeating, fallacious arguments providing thus a simple and semantic way to identify such arguments. We have seen that in particular this allows us to semantically capture proof by contradiction in formal logical systems which helps to form a simple correspondence between formal reasoning and argumentation. On the other hand it points the way to how to include such principles of reasoning in informal reasoning.

Based on the above elements we have a framework of argumentation that can encompass formal and informal logic in a uniform way. This is not to say, though, that we can capture within the formalism the full extent of informal logic with all its challenging nuances. Nevertheless, it opens up a way of studying argumentation in this larger context, which has recently been called *cognitive argumentation*, to distinguish it from abstract argumentation in AI and to explicate the emphasis on the informal human reasoning. This approach aims to address the challenges of informal logic by guiding and informing the development of computational frameworks of argumentation, through empirical findings and models of cognition from studies in cognitive psychology and cognitive science. General cognitive principles drawn from independent observations of human reasoning and thinking behavior can inform both the construction of the argumentation framework and importantly the computational dialectic process of argumentation and ensuing reasoning forms.

Within such a framework of cognitive argumentation the major tasks are to choose cognitively valid argument schemes and to address the need for contextual reasoning by, typically extra-logical, processes of selecting the relevant arguments to reason with and, in particular, to use context-based preferences between schemes that can dynamically change with changes in the current environment. This provides a way to encompass the required adaptability and flexibility of informal reasoning. Early studies of this cognitive argumentation approach on human syllogistic reasoning and human conditional reasoning can be found in (Saldanha and Kakas 2018 and 2019). How we can dynamically vary the degree of flexibility in the argumentation framework is a matter of further investigation and of

importance for the deeper understanding of human reasoning. For example, as observed in human syllogistic and conditional reasoning, the human population exhibits a degree of variation in their reasoning and we are challenged to find further principles and heuristics of argumentation that would help us understand this variability.

References

Baroni, P., D. Gabbay and M. Giacomin. 2018. *Handbook of formal argumentation*. College Publications.

Bench-Capon, T. J. M. and P. E. Dunne. 2007. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15): 619–641.

Besnard, P. and A. Hunter. 2008. *Elements of argumentation*. MIT Press.

Bondarenko, A., F. Toni and R. Kowalski. 1993. An assumption-based framework for non-monotonic reasoning. In *Proceedings of the 2nd international workshop on logic programming and non-monotonic reasoning (LPNMR)*, 171–189.

Diakidoy, I.-A., A. Kakas, L. Michael and R. Miller. 2014. Story comprehension through argumentation. In *Proceedings of the 5th international conference on computational models of argument (COMMA2014)*, 31–42.

Diakidoy, I.-A., A. Kakas, L. Michael and R. Miller. 2015. STAR: A system of argumentation for story comprehension and beyond. In *Proceedings of the 12th international symposium on logical formalizations of commonsense reasoning (Commonsense2015)*.

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

van Eemeren, F. H., B. Garssen, E. C. W. Krabbe, A. F. S. Henkemans, B. Verheij and J. H. M. Wagemans, eds. 2014. *Handbook of argumentation theory*. Springer Reference.

Evans, J. S. B. T., S. E. Newstead and R. M. J. Byrne. 1993. *Human reasoning: The psychology of deduction*. Lawrence Erlbaum Associates.

Gentzen, G. 1935. Unstersuchungenuber das Logische Schliessen. *Mathematische Zeitschrift,* 39: 176–210. English translation: M. Szabo, ed. 1969. *The collected papers of Gehard Gentzen*. Amsterdam: North Holland.

Gordon, T. F. and D. Walton. 2006. The Carneades argumentation framework—using presumptions and exceptions to model critical questions. In *Computational models of argument: Proceedings of COMMA 2006*, 195–207.

Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics: Vol. 3: Speech acts*, eds. P. Cole and J. L. Morgan, 41–58. New York: Academic Press.

Groarke, L. 2017. Informal logic. In *The Stanford encyclopaedia of philosophy* (spring 2017 edition)*,* ed. E. N. Zalta. Metaphysics Research Lab, Stanford University.

Johnson, R. H. and J. A. Blair. 2000. Informal logic: An overview. *Informal Logic*, 20: 94–108.

Johnson-Laird, P. 1983. *Mental models*. Cambridge University Press.

Kakas, A. C., F. Toni and P. Mancarella. 2014. Argumentation logic. In *Proceedings of the 5th international conference on computational models of argument (COMMA2014)*, 12–27.

Kakas, A. C., P. Mancarella and F. Toni. 2018. On argumentation logic and propositional logic. *Studia Logica*, 106(2): 237–279.

Khemlani, S. and P. N. Johnson-Laird. 2012. Theories of the syllogism: A meta- analysis. *Psychological Bulletin*, 138: 427–457.

Kowalski, R. 2011. *Computational logic and human thinking: How to be artificially intelligent.* New York, NY, USA: Cambridge University Press.

Lorenz, K. 2001. Basic objectives of dialogue logic in historical perspective. *Synthese*, 127(1): 255–263.

Modgil, S. and H. Prakken. 2013. A general account of argumentation with preferences. *Artificial Intelligence*, 195: 361–397.

Paglieri, F. 2013. Argumentation, decision and rationality. In *OSSA Conference Archive*.

Pollock, J. L. 1987. Defeasible reasoning. *Cognitive Science*, 11(4): 481–518.

Pollock, J. L. 1991. Self-defeating arguments. *Minds and Machines*, 1: 367–392.

Pollock, J. L. 1995. *Cognitive carpentry: A blueprint for how to build a person*. Studies in Argumentation Theory.

Prakken, H. 2005. AI and law, logic and argument schemes. *Argumentation*, 19(3): 303–320.

Rahwan, I. and G. R. Simari. 2009. *Argumentation in artificial intelligence*. Springer Publishing Company.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence*, 13(1,2): 81–132.

Ruckert, H. 2001. Why dialogical logic? In *Essays on non-classical logic (Advances in Logic 1),* ed. H. Wansing, 165–185. River Edge: World Scientific.

Saldanha, E.-A. D. and A. C. Kakas. 2018. Cognitive argumentation for human syllogistic reasoning. Under review.

Saldanha, E.-A. D. and A. C. Kakas. 2019. The suppression task in cognitive argumentation. Under review.

Stenning, K. and M. van Lambalgen. 2008a. *Human reasoning and cognitive science*. MIT Press.

Thaler, R. H. 1991. Quasi rational economics. New York, NY, US: Russell Sage Foundation.

Toulmin, S. 1958. *The uses of argument*. Cambridge University Press.

Tversky, A. and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157): 1124–1131.

Walton, D. 1995. *Argumentation schemes for presumptive reasoning*. Bradford Books.

Walton, D. 2008. *Informal logic: A pragmatic approach*. Cambridge University Press.

Walton, D. and T. F. Gordon. 2015. Formalizing informal logic. *Informal Logic*, 35(4): 508–538.