

# On Natural Language Generation of Formal Argumentation

Federico Cerutti<sup>1</sup>, Alice Toniolo<sup>2</sup>, and Timothy J. Norman<sup>3</sup>

<sup>1</sup> Università degli Studi di Brescia  
Dipartimento di Ingegneria dell'Informazione  
Brescia, Italy

<sup>2</sup> University of St. Andrews  
School of Computer Science  
St. Andrews, UK

<sup>3</sup> University of Southampton  
Department of Electronics and Computer Science  
Southampton, UK

**Abstract.** In this paper we provide a first analysis of the research questions that arise when dealing with the problem of communicating pieces of formal argumentation through natural language interfaces. It is a generally held opinion that formal models of argumentation naturally capture human argument, and some preliminary studies have focused on justifying this view. Unfortunately, the results are not only inconclusive, but seem to suggest that explaining formal argumentation to humans is a rather articulated task. Graphical models for expressing argumentation-based reasoning are appealing, but often humans require significant training to use these tools effectively. We claim that natural language interfaces to formal argumentation systems offer a real alternative, and may be the way forward for systems that capture human argument.

## 1 Introduction

Our aim here is to explore the challenges that we need to face when thinking about natural language interfaces to formal argumentation. Dung [12] states that formal argumentation “captures naturally the way humans argue to justify their solutions to many social problems.” This is one of the most common claims used to support research in formal argumentation. More recently there have been a number of empirical studies to investigate this claim [24, 10, 26, 27]. The results, however, have been far from conclusive.

The use of graphical models to represent arguments is the most common approach used in the formal argumentation community to capture argument structures. This has been successfully applied in a number of real world situations: Laronge [19] (an American barrister and researcher), for example, describes how he used argumentation graphs during trials. Despite this, our claim is that to produce and to consume a graphical representation of a structure of arguments there is need for significant levels of training.

Instead of training users on another (graphical) language for representing argument structures, we can leverage our societal model, through which we are trained in reading

and writing; that is, using natural language. We claim that natural language representations of formal argumentation are the way forward to develop formal models that capture human argumentation. In this paper we investigate one aspect of natural language interfaces to formal argumentation: moving from formal arguments to natural language text by exploiting Natural Language Generation (NLG) systems.

In CISpaces.org [9, 8] we followed a rather pragmatic approach. Indeed, we implemented: (1) a template-based NLG system; (2) a greedy, heuristics-based approach for chaining together premises and conclusion of arguments; (3) an assert-justify writing style suitable for speed reading.

In this paper, we go beyond the current simple template-based NLG system implemented in CISpaces.org, and we ground our investigation on an existing example (Section 2) from collaboration between the BBC and the Dundee argumentation group [20], namely an excerpt of the BBC Radio 4 Moral Maze programme from 22nd June 2012. In this way the reader can always relate to the original piece of text from where our investigation started. The excerpt has been already formalised into an argument network (i.e. a graph linking together different pieces of information together to display the web of arguments exchanged). In Figure 1 and in Section 3 we review all the necessary elements for our investigation: the notion of argumentation schemes; how to represent argument networks in the Argument Interchange Format (AIF); a (simple) approach to structured argumentation to build arguments and approximate arguments from an AIF argument network; Dung’s theory of argumentation; and basic elements of NLG.

In NLG one of the most difficult tasks is to determine the communicative goal; i.e., deciding what we would like to communicate. Therefore, Section 4 is entirely dedicated to investigating relevant communicative goals in the context of formal argumentation.

The result we wish to achieve in this paper is a blueprint that outlines the complex research questions and their dense interconnections that our community needs to address in order to identify models that naturally capture human argumentation.

## 2 Running Example

On 22nd June 2012, in the middle of the European debt crisis, the Moral Maze program on BBC Radio 4 addressed the topic of individual and national debt. Among others, Nick Dearden, director of the Jubilee Debt Campaign, and Claire Fox, from the Institute of Ideas, were ‘witnesses’ (contributors offering a specific point of view in the debate) during the program.

What follows is an excerpt that has been analysed to identify the argument network in the dialogue and made available at <http://aifdb.org/argview/1724> [20]. In this paper we focus on the sub-part of the argument network depicted in Figure 1 and we highlight in the text the elements that contribute to the generation of such an argument network.

CLAIRE FOX: I understand that. I suppose my concern is just this: I want the freedom to be able to write off debts but — I’m sure you recognise this — there is this sort of sense amongst a lot of young people, who just think, “I want that, so I’ll have that now. Thank you.” And so, if you want the moral hazard, instead of kind of just going on about the bankers, is there not a danger that if we

just said we'd write off debt, that it actually isn't very helpful for our side, for ordinary people, to actually have that? [T1] There's no discipline there [T2]. In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt?" [T3]

NICK DEARDEN: In some ways I agree with you. If you want the economy to run smoothly, you have to incentivise certain types of behaviour. [T4] So, for example in South Korea, in terms of how South Korean grew, it did incentivise saving, at certain times, by certain economic policies [T5]. On the other hand, I think what people don't realise, or only half realise, is the fact that we have actually written off massive amounts of debt [T6]. But it certainly isn't the debts of the people who most need it in society [T7].

### 3 Background

#### 3.1 Argumentation Schemes

Argumentation schemes [33, 34] are abstract reasoning patterns commonly used in everyday conversational argumentation, legal, scientific argumentation, etc. Schemes have been derived from empirical studies of human argument and debate. They can capture traditional deductive and inductive approaches as well as plausible reasoning. Each scheme has a set of critical questions that represents standard ways of critically probing into an argument to find aspects of it that are open to criticism.

For instance, in the dialogue reported in Section 2, part of Nick's position can be mapped into an argument from example that has the following structure:

Premise: In this particular case, individual  $a$  has property  $F$  and also property  $G$ .

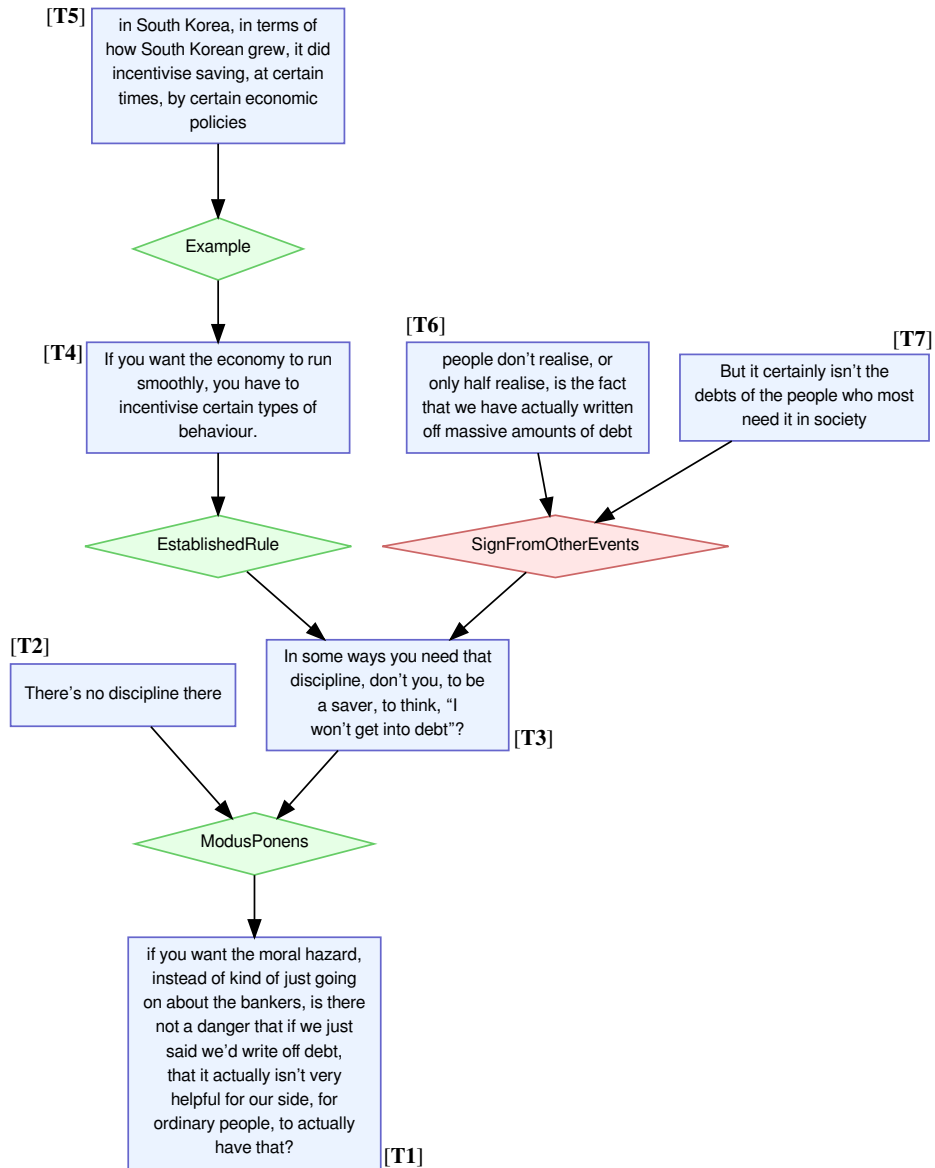
Conclusion: Therefore, generally, if  $x$  has property  $F$ , then it also has property  $G$ .

One of the critical questions here is: "Is the proposition claimed in the premise in fact true?"

#### 3.2 Representing an Argument Network

The Argument Interchange Format (AIF) [11, 24] is the current proposal for a standard notation for argument structures. It is based on a graph that specifies two types of nodes: information nodes (or I-nodes) and scheme nodes (or S-nodes). These are represented by two disjoint sets,  $\mathcal{N}_I \cup \mathcal{N}_S = \mathcal{N}$  and  $\mathcal{N}_I \cap \mathcal{N}_S = \emptyset$ , where information nodes represent claims, premises, data, etc., and scheme nodes capture the application of patterns of reasoning belonging to a set  $\mathcal{S} = \mathcal{S}^R \cup \mathcal{S}^C \cup \mathcal{S}^P$ ,  $\mathcal{S}^R \cap \mathcal{S}^C = \mathcal{S}^C \cap \mathcal{S}^P = \mathcal{S}^P \cap \mathcal{S}^R = \emptyset$ . Reasoning patterns can be of three types: rule of inference  $\mathcal{S}^R$ ; criteria of preference  $\mathcal{S}^P$ ; and criteria of conflicts  $\mathcal{S}^C$ .

The relation  $\text{fulfils} \subseteq \mathcal{N}_S \times \mathcal{S}$  expresses that a scheme node instantiates a particular scheme. Scheme nodes, moreover, can be one of three types: rule of inference application nodes  $\mathcal{N}_S^{RA}$ ; preference application nodes  $\mathcal{N}_S^{PA}$ ; or conflict application nodes  $\mathcal{N}_S^{CA}$ , with  $\mathcal{S} = \mathcal{N}_S^{RA} \cup \mathcal{N}_S^{PA} \cup \mathcal{N}_S^{CA}$ , and  $\mathcal{N}_S^{RA} \cap \mathcal{N}_S^{PA} = \mathcal{N}_S^{PA} \cap \mathcal{N}_S^{CA} = \mathcal{N}_S^{CA} \cap \mathcal{N}_S^{RA} = \emptyset$ .



**Fig. 1.** Argument network for the running example, as obtained from <http://aifdb.org/argview/1724> (edited version)

Nodes are connected by edges whose semantics is implicitly defined by their use. For instance, an information node connected to a RA scheme node, with the arrow terminating in the latter, would suggest that the information node serves as a premise for an inference rule. Figure 1 shows an AIF representation of the arguments exchanged in the dialogue introduced in Section 2. Rectangular nodes represent information nodes, while rhombic ones represent scheme nodes: green for RA nodes, and red for CA nodes.

### 3.3 Deductive Argumentation

Using deductive argumentation means that each argument is defined using a logic, and in the following we adopt the simple, but elegant logic proposed by Besnard & Hunter [3]. Thus, we let  $\mathcal{L}$  be a logical language. If  $\alpha$  is an atom in  $\mathcal{L}$ , then  $\alpha$  is a *positive literal* in  $\mathcal{L}$ , and  $\neg\alpha$  is a *negative literal* in  $\mathcal{L}$ . For a literal  $\beta$ , the *complement* of the positive literal  $\beta = \alpha$  is  $\bar{\beta} = \neg\alpha$  (resp. if  $\beta = \neg\alpha$  is not a positive literal, its complement is the positive literal  $\bar{\beta} = \alpha$ ).

A *simple rule* is of the form  $\alpha_1 \wedge \dots \wedge \alpha_k \rightarrow \beta$  where  $\alpha_1, \dots, \alpha_k, \beta$  are literals. A *simple logical knowledge base* is a set of literals and a set of simple rules. Given a simple logic knowledge base,  $\Delta$ , the *simple consequence relation*  $\vdash_s$  is defined, such that  $\Delta \vdash_s \beta$  if and only if there is a rule  $\alpha_1 \wedge \dots \wedge \alpha_n \rightarrow \beta \in \Delta$  and  $\forall i$  either  $\alpha_i \in \Delta$  or  $\Delta \vdash_s \alpha_i$ . Now, given  $\Phi \subseteq \Delta$  and a literal  $\alpha$ ,  $\langle \Delta, \alpha \rangle$  is a *simple argument* if and only if  $\Phi \vdash_s \alpha$  and  $\nexists \Phi' \subsetneq \Phi$  such that  $\Phi' \vdash_s \alpha$ .  $\Phi$  is the support (or premises, assumptions) of the argument, and  $\alpha$  is the claim (or conclusion) of the argument. Given an argument  $\mathbf{a} = \langle \Phi, \alpha \rangle$ , the function  $\text{Support}(\mathbf{a})$  returns  $\Phi$ , and  $\text{Claim}(\mathbf{a})$  returns  $\alpha$ .

For simple arguments  $\mathbf{a}$  and  $\mathbf{b}$  we consider the following types of simple attack:

- $\mathbf{a}$  is a *simple undercut* of  $\mathbf{b}$  if there is a simple rule  $\alpha_1 \wedge \alpha_k \rightarrow \beta$  in  $\text{Support}(\mathbf{b})$  and there is an  $\alpha_i \in \{\alpha_1, \dots, \alpha_k\}$  such that  $\text{Claim}(\mathbf{a})$  is the complement of  $\alpha_i$ ;
- $\mathbf{a}$  is a *simple rebut* of  $\mathbf{b}$  if  $\text{Claim}(\mathbf{a})$  is the complement of  $\text{Claim}(\mathbf{b})$ .

Following Black & Hunter [4], an *approximate argument* is a pair  $\langle \Phi, \alpha \rangle$ . If  $\Phi \vdash_s \alpha$ , then  $\langle \Phi, \alpha \rangle$  is also *valid*; if  $\Phi \not\vdash_s \perp$ , then  $\langle \Phi, \alpha \rangle$  is also *consistent*; if  $\Phi \vdash_s \alpha$ , and there is no  $\Phi' \subsetneq \Phi$  such that  $\Phi' \vdash_s \alpha$ , then  $\langle \Phi, \alpha \rangle$  is also *minimal*; if  $\Phi \vdash_s \alpha$ , and  $\Phi \not\vdash_s \perp$ , then  $\langle \Phi, \alpha \rangle$  is also *expansive* (i.e. it is valid and consistent, but it may have unnecessary premises).

Building on top of Figure 1 and transforming each  $\mathcal{N}_S^{RA}$  node into a simple rule, a simple knowledge base for our running example is:

$$\Delta_m = \left\{ \begin{array}{l} [\mathbf{T1}], [\mathbf{T2}], [\mathbf{T3}], [\mathbf{T4}], [\mathbf{T5}], \\ [\mathbf{T5}] \rightarrow [\mathbf{T4}], \\ [\mathbf{T4}] \rightarrow [\mathbf{T3}], \\ [\mathbf{T2}] \wedge [\mathbf{T3}] \rightarrow [\mathbf{T1}], \\ [\mathbf{T6}] \wedge [\mathbf{T7}] \rightarrow [\mathbf{T3}] \end{array} \right\}$$

Therefore, the following are the simple arguments that can be built from  $\Delta_m$ :

$$\mathcal{A}_m = \left\{ \begin{array}{l} \mathbf{a} = \langle \{[\mathbf{T5}], [\mathbf{T5}] \rightarrow [\mathbf{T4}], [\mathbf{T4}]\}, \\ \mathbf{b} = \langle \{[\mathbf{T4}], [\mathbf{T4}] \rightarrow [\mathbf{T3}]\}, [\mathbf{T3}] \rangle, \\ \mathbf{c} = \langle \{[\mathbf{T2}], [\mathbf{T3}], [\mathbf{T2}] \wedge [\mathbf{T3}] \rightarrow [\mathbf{T1}]\}, [\mathbf{T1}] \rangle, \\ \mathbf{d} = \langle \{[\mathbf{T6}], [\mathbf{T7}], [\mathbf{T6}] \wedge [\mathbf{T7}] \rightarrow [\mathbf{T3}]\}, [\mathbf{T3}] \rangle \end{array} \right\}$$

with  $\mathbf{d}$  rebutting  $\mathbf{b}$ , and  $\mathbf{d}$  undercutting  $\mathbf{c}$ .

However, there are many more approximate arguments. For instance,  $\mathbf{c}' = \langle \{[\mathbf{T5}], [\mathbf{T4}], [\mathbf{T3}], [\mathbf{T2}], [\mathbf{T5}] \rightarrow [\mathbf{T4}], [\mathbf{T4}] \rightarrow [\mathbf{T3}], [\mathbf{T2}] \wedge [\mathbf{T3}] \rightarrow [\mathbf{T1}]\}, [\mathbf{T1}] \rangle$  is an approximate argument in favour of  $[\mathbf{T1}]$  taking into consideration all the inferences that might help concluding it. Conversely,  $\mathbf{c}'' = \langle \{\}, [\mathbf{T1}] \rangle$  is the minimal (invalid) approximate argument in favour of  $[\mathbf{T1}]$ .

### 3.4 Abstract Argumentation

An argumentation framework [12] consists of a set of arguments and a binary attack relation between them.

**Definition 1.** An argumentation framework (AF) is a pair  $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$  where  $\mathcal{A}$  is a set of arguments and  $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ . We say that  $\mathbf{b}$  attacks  $\mathbf{a}$  iff  $\langle \mathbf{b}, \mathbf{a} \rangle \in \mathcal{R}$ , also denoted as  $\mathbf{b} \rightarrow \mathbf{a}$ . The set of attackers of an argument  $\mathbf{a}$  will be denoted as  $\mathbf{a}^- \triangleq \{\mathbf{b} : \mathbf{b} \rightarrow \mathbf{a}\}$ , the set of arguments attacked by  $\mathbf{a}$  will be denoted as  $\mathbf{a}^+ \triangleq \{\mathbf{b} : \mathbf{a} \rightarrow \mathbf{b}\}$ . We also extend these notations to sets of arguments, i.e. given  $E, S \subseteq \mathcal{A}$ ,  $E \rightarrow \mathbf{a}$  iff  $\exists \mathbf{b} \in E$  s.t.  $\mathbf{b} \rightarrow \mathbf{a}$ ;  $\mathbf{a} \rightarrow E$  iff  $\exists \mathbf{b} \in E$  s.t.  $\mathbf{a} \rightarrow \mathbf{b}$ ;  $E \rightarrow S$  iff  $\exists \mathbf{b} \in E, \mathbf{a} \in S$  s.t.  $\mathbf{b} \rightarrow \mathbf{a}$ ;  $E^- \triangleq \{\mathbf{b} \mid \exists \mathbf{a} \in E, \mathbf{b} \rightarrow \mathbf{a}\}$  and  $E^+ \triangleq \{\mathbf{b} \mid \exists \mathbf{a} \in E, \mathbf{a} \rightarrow \mathbf{b}\}$ .

Each argumentation framework, therefore, has an associated directed graph where the vertices are the arguments, and the edges are the attacks.

The basic properties of conflict-freeness, acceptability, and admissibility of a set of arguments are fundamental for the definition of argumentation semantics.

**Definition 2.** Given an AF  $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$ :

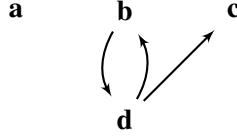
- a set  $S \subseteq \mathcal{A}$  is a conflict-free set of  $\Gamma$  if  $\nexists \mathbf{a}, \mathbf{b} \in S$  s.t.  $\mathbf{a} \rightarrow \mathbf{b}$ ;
- an argument  $\mathbf{a} \in \mathcal{A}$  is acceptable with respect to a set  $S \subseteq \mathcal{A}$  of  $\Gamma$  if  $\forall \mathbf{b} \in \mathcal{A}$  s.t.  $\mathbf{b} \rightarrow \mathbf{a}$ ,  $\exists \mathbf{c} \in S$  s.t.  $\mathbf{c} \rightarrow \mathbf{b}$ ;
- a set  $S \subseteq \mathcal{A}$  is an admissible set of  $\Gamma$  if  $S$  is a conflict-free set of  $\Gamma$  and every element of  $S$  is acceptable with respect to  $S$ , i.e.  $S \subseteq \mathcal{F}_\Gamma(S)$ .

An argumentation semantics  $\sigma$  prescribes for any AF  $\Gamma$  a set of extensions, denoted as  $\mathcal{E}_\sigma(\Gamma)$ , namely a set of sets of arguments satisfying the conditions dictated by  $\sigma$ . For instance, here is the definition of preferred (denoted as PR) semantics.

**Definition 3.** Given an AF  $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$ , a set  $S \subseteq \mathcal{A}$  is a preferred extension of  $\Gamma$ , i.e.  $S \in \mathcal{E}_{\text{PR}}(\Gamma)$ , iff  $S$  is a maximal (w.r.t. set inclusion) admissible set of  $\Gamma$ .

Given a semantics  $\sigma$ , an argument  $\mathbf{a}$  is said to be *credulously accepted* w.r.t.  $\sigma$  if  $\mathbf{a}$  belongs to at least one  $\sigma$ -extension.  $\mathbf{a}$  is *skeptically accepted* w.r.t.  $\sigma$  if  $\mathbf{a}$  belongs to all the  $\sigma$ -extensions.

It can be noted that each complete extension  $S$  implicitly defines a three-valued labelling function  $\text{Lab}$  on the set of arguments: an argument  $\mathbf{a}$  is labelled in iff  $\mathbf{a} \in S$ ; is labelled out iff  $\exists \mathbf{b} \in S$  s.t.  $\mathbf{b} \rightarrow \mathbf{a}$ ; and is labelled undec if neither of the above conditions holds. In the light of this correspondence, argumentation semantics can be equivalently defined in terms of labellings rather than of extensions [7, 2].



**Fig. 2.** The AF  $\Gamma_m$  for Figure 1 interpreted using deductive argumentation.

We can now introduce the concept of the *issues* of an argumentation framework whose status is enough to determine the status of all the arguments in the framework [14, 5, 6].

**Definition 4.** Given an AF  $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$  and  $\mathcal{L}$  the set of all complete labellings of  $\Gamma$ , for any  $\mathbf{a}, \mathbf{b} \in \mathcal{A}$ ,  $\mathbf{a} \equiv \mathbf{b}$  iff  $\forall \mathcal{L}ab \in \mathcal{L}$ ,  $\mathcal{L}ab(\mathbf{a}) = \mathcal{L}ab(\mathbf{b})$ ; or  $\forall \mathcal{L}ab \in \mathcal{L}$ ,  $(\mathcal{L}ab(\mathbf{a}) = \text{in} \iff \mathcal{L}ab(\mathbf{b}) = \text{out}) \wedge (\mathcal{L}ab(\mathbf{a}) = \text{out} \iff \mathcal{L}ab(\mathbf{b}) = \text{in})$ .

The set of arguments in the equivalent class  $\equiv$  is the set of issues of  $\Gamma$ :

$$\mathcal{E}_{\text{issues}}(\Gamma) = \left\{ S \subseteq \mathcal{A} \mid \forall \langle \mathbf{a}, \mathbf{b} \rangle \in S \times S, \mathbf{a} \equiv \mathbf{b}; \text{ and } \forall S' \supset S, \exists \langle \mathbf{c}, \mathbf{d} \rangle \in S' \times S', \neg(\mathbf{c} \equiv \mathbf{d}) \right\}$$

Continuing with our running example, Figure 2 depicts the argumentation framework from Section 3.3 applying deductive argumentation on the argument network of Figure 1:  $\Gamma_m = \langle \mathcal{A}_m, \mathcal{R}_m \rangle = \langle \{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}, \{\mathbf{d} \rightarrow \mathbf{b}, \mathbf{b} \rightarrow \mathbf{d}, \mathbf{d} \rightarrow \mathbf{c}\} \rangle$ .

There are two preferred extensions:  $\mathcal{E}_{\text{PR}}(\Gamma_m) = \{\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}, \{\mathbf{a}, \mathbf{d}\}\}$ . Moreover,  $\{\mathbf{b}, \mathbf{d}\} \in \mathcal{E}_{\text{issues}}(\Gamma_m)$ , and if  $\mathcal{L}ab(\mathbf{b}) = \text{in}$ ,  $\mathcal{L}ab(\mathbf{d}) = \text{out}$  (and viceversa).

### 3.5 Natural Language Generation

A Natural Language Generation (NLG) system requires [25]:

- A knowledge source to be used;
- A communicative goal to be achieved;
- A user model; and
- A discourse history.

In general, the knowledge source is the information about the domain, while the communicative goal describes the purpose of the text to be generated. The user model is a characterisation of the intended audience, and the discourse history is a model of what has been said so far.

An NLG system divides processing into a pipeline [25] composed of the three stages described in Table 1. First it determines the content and structure of a document (*document planning*); then it looks at syntactic structures and choices of words that needs to be used to communicate the content chosen in the document planning (*microplanning*). Finally, it maps the output of the microplanner into text (*realisation*).

Each stage includes tasks that can be primarily concerned with either content or structure. Document planning requires content determination — deciding what information should be communicated in the output document — and document structuring — how to order the information to be conveyed.

**Table 1.** Modules and tasks of a NLG system, adapted from [25, Figure 3.1].

<b>Module</b>	<b>Content task</b>	<b>Structure task</b>
Document planning	Content determination	Document structuring
Microplanning	Lexicalisation Referring expression generation	Aggregation
Realisation	Linguistic realisation	Structure realisation

Microplanning requires (1) lexicalisation — deciding what syntactic constructions our NLG system should use; (2) referring expression generation — how to relate with entities; and (3) aggregation — how to map the structures created by the document planning onto linguistic structures such as sentences and paragraphs.

Document planning and microplanning are the most strategic and complex modules in this pipeline. They focus on identifying the communicative goal and how it relates to the user model, thus producing the blueprint of the document that will be generated.

It is the responsibility of the document planning module, in particular of the document structuring task, to consider the rhetorical relations (or discourse relations) that hold between messages or groups of messages. For instance, Rhetorical Structure Theory (RST) [21] stresses the idea that the coherency of a text depends on the relationships between pairs of *text spans* (an uninterrupted linear interval of text) a *nucleus* and a *satellite*. In [21] a variety of relationships are provided, but for the purpose of this work, we focus on:

**Evidence** The nucleus contains a claim, while the satellite(s) contain(s) evidence supporting such a claim.

**Justify** The nucleus contains a claim, while the satellite(s) contain(s) justification for such a claim.

**Antithesis** Nucleus and satellite are in contrast.

Finally, once the more strategical tasks are performed, there is need for linguistic realisation — from abstract representations of sentences to actual text — and structure realisation — converting abstract structures such as paragraphs and sections into the mark-up symbols chosen for the document. The realisation module is the most algorithmic in this pipeline, and there are already several implementations for supporting it, for instance SimpleNLG [15].

In the next section we highlight the cases of document planning and microplanning we believe are most interesting from an argumentation perspective.

## 4 Generating Natural Language Interfaces to Formal Argumentation

Let us now return to our running example to illustrate four relevant communicative goals. For the moment we are not making any assumptions about the user model, and



we will assume no pre-existing discourse history. Therefore, given that the knowledge source is fixed, we envisage the following four communicative goals, each of which raises interesting and challenging questions:

1. Presenting a single argument or an approximate argument;
2. Presenting an entire argument network;
3. Explaining the acceptability status of a single argument or an approximate argument; and
4. Explaining the extensions, given some semantics.

In the following, we discuss elements of content determination for each of these goals; i.e. deciding what messages should be included in the document to be generated. Examples will be provided based on our running example. In parts, the generated texts will sound a little awkward because we deliberately chose not to modify the content of the arguments. We will elaborate on this in Section 5.

#### 4.1 Presenting a Simple or an Approximate Argument

Simple and approximate arguments are composed of premises and a claim. There are two main strategies traditionally adopted to represent such a construct:

- *forward writing*, from premises to claim;
- *backward writing*,<sup>4</sup> from claim to premises.

Let us consider the argument  $\mathbf{b} = \langle \{[\mathbf{T4}], [\mathbf{T4}] \rightarrow [\mathbf{T3}], [\mathbf{T3}]\rangle$ . In the case of *forward writing*, we can write something like:

If you want the economy to run smoothly, you have to incentivise certain type of behaviour. [T4] In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. [T3]

An explicit signal such as Therefore can be used to highlight the *Justify* relation:

If you want the economy to run smoothly, you have to incentivise certain type of behaviour. [T4]

Therefore In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. [T3]

Similarly in the case of *backward writing*, we may write:

In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. [T3]

Indeed If you want the economy to run smoothly, you have to incentivise certain type of behaviour. [T4]

---

<sup>4</sup>Often also named *assert-justify*.

More interesting is the case of considering an approximate argument. Assuming that the communicative goal here is not to confuse the reader, it is probably reasonable not to include irrelevant elements in the presentation. Relevance theory may be seen as an attempt to analyse an essential feature of most human communication: the expression and recognition of intentions [17, 29].

For instance, the First or Cognitive Principle of Relevance proposed by Sperber and Wilson [28, 29] states: *Human cognition tends to be geared to the maximisation of relevance*. Therefore, it would be evidence of poor judgment to expand the argument by considering *irrelevant* elements. Unfortunately, defining the concept of *relevance* in formal argumentation is a non-trivial task [32, 23]. To take a reasonable starting point, let us adapt the definition of relevant evidence from the Rule 401 of the Federal Rules of Evidence<sup>5</sup> to our context. Therefore, an approximate argument is relevant for an argument if:

1. Its premises have any tendency to make the argument's conclusion more or less probable than it would be without them; and
2. It provides additional information that might advance the debate.

Given this (loose) definition of relevance, we could argue that the approximate argument  $\mathbf{b}' = \langle \{[\mathbf{T5}], [\mathbf{T5}] \rightarrow [\mathbf{T4}], [\mathbf{T4}] \rightarrow [\mathbf{T3}]\}, [\mathbf{T3}] \rangle$  is relevant for  $\mathbf{b}$ . Similarly as before, there are different possibilities to write such a chain of inferences. In addition, it raises questions related to microplanning: perhaps, we desire to merge together different pieces such as in the following example using *backward writing* style.

In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. [T3]

Indeed [T4] If you want the economy to run smoothly, you have to incentivise certain type of behaviour. [T4] , e.g. in South Korea, in terms of how South Korean grew, it did incentivise saving, at certain times, by certain economic policies [T5]

[T4] and [T5] are merged by a comma and [e.g.]. In this way, we highlight the *Evidence* relation and identify the connection as an argument from example instead of forming two independent sentences, which is an aspect traditionally related to microplanning.

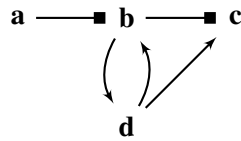
Finally, if we have knowledge that an inference rule fulfils an argumentation scheme, then we can also map the elements of an argument into such a scheme. For instance, argument  $\mathbf{b}$  is classified as an argument by established rule, that is represented by the following scheme:

Major Premise: If carrying out types of actions including the state of affairs A is the established rule for x, then (unless the case is an exception), x must carry out A.

Minor Premise: Carrying out types of actions including state of affairs A is the established rule for a.

Conclusion: Therefore a must carry out A.

<sup>5</sup>[https://www.law.cornell.edu/rules/fre/rule\\_401](https://www.law.cornell.edu/rules/fre/rule_401) (on 13/05/2017)



**Fig. 3.** The  $AF \Gamma_m$  for the argument network of Figure 1 interpreted using deductive argumentation in a bipolar-inspired framework, where the squared arrows represents generic dependencies based on inferences.

It is interesting to note that the minor premise is left implicit in the formalisation of our example, and it might be an element that should be highlighted to the user. How to report an implicit premise may depend on the communicative goal. For example, if the system is intended for users to improve and make their arguments more explicit we may generate the text:

In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. [T3] Indeed If you want the economy to run smoothly, you have to incentivise certain type of behaviour. [T4]  
although we have no evidence that this is the established rule.

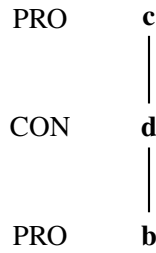
In this case, the fact that there is a lack of the minor premise is added to the generated text. On the other hand, if the system is intended to report an analysis of a conversation, we need to take into account that the premise is implicit as it may be already known by the participants. Hence, the system could assume that the premise holds, and generate an explicit sentence and we assume that this is the established rule. Further research is, however, needed in understanding how to generate text in case of unstated premises (e.g., in case of enthymemes [33, 4]). With similar considerations, we could also include critical questions that have either already been answered, or that, if answered, could strengthen the argument.

#### 4.2 Presenting an Argument Network

One of the strengths of formal argumentation is its ability to handle conflicts. In the previous section we focused on how to represent a single argument. However, even in our small running example, we can note that (1) there are more than one argument; and (2) there are conflicts among them (Figures 1 and 2).

The simplest strategy for representing an argument network is just to enumerate all the arguments and to list the conflicts among them, eventually linking that with critical questions if the information is available (e.g. This counterargument answers the critical question that states . . .).

However, we argue that there is a better approach, motivated by the need to satisfy Grice's maxim of relation and relevance [16]; i.e. that in a conversation one needs to be relevant and say things that are pertinent to the discussion. Let us consider Figure



**Fig. 4.** Dispute tree to prove that **c** is credulously accepted w.r.t. preferred semantics in the  $AF \Gamma_m$  for the argument network of Figure 1 interpreted using deductive argumentation.

3 that depicts the  $\Gamma_m$  argumentation framework (cf. Fig. 2) annotated with a second, bipolar-inspired [1],<sup>6</sup> binary relation between arguments. Such a relation is grounded in the original argument network, Figure 1.

Once such an annotated graph is obtained, a line of reasoning annotated together (e.g.  $\langle \mathbf{a}, \mathbf{b}, \mathbf{c} \rangle$ ) in Figure 3) could be merged together in order to provide a single approximate argument. Once all the lines of reasoning annotated together are identified and merged, then we need a strategy to order their presentation; e.g. ordering by length of each line of reasoning versus ordering by the number of attacks received. We also need appropriate signals for *antithesis* relations need to be used. For instance:

if you want the moral hazard, instead of kind of just going on about the bankers, is there not a danger that if we just said we'd write off debt, that it actually isn't very helpful for our side, for ordinary people, to actually have that?  
**[T1]** Indeed There's no discipline there **[T2]** and In some ways you need that discipline, don't you, to be a saver, to think, "I won't get into debt"?. **[T3]**  
**Indeed** If you want the economy to run smoothly, you have to incentivise certain type of behaviour. **[T4]** , e.g. in South Korea, in terms of how South Korean grew, it did incentivise saving, at certain times, by certain economic policies **[T5]** However, people don't realise, or only half realise, is the fact that we have actually written off massive amounts of debt **[T6]** and But it certainly isn't the debts of the people who most need it in society **[T7]**.

### 4.3 Explaining the Acceptability Status of a Single Argument or an Approximate Argument

Two of the traditional decision problems in abstract argumentation [13] are checking whether an argument is credulously or skeptically accepted (or not). An interested user might select one of the arguments (or one of the propositions in the argument network) and ask whether or not it is credulously or skeptically accepted.

<sup>6</sup>By *bipolar-inspired* we intend that such a relation does not represent conflicts. At the same time, however, we are not in the position to argue that it would be a relationship of *support*.

To answer such a query, we can exploit research in proof theory and argument games [22]. For instance, to prove that **c** (in Figure 2) is credulously accepted according to preferred semantics, we can (1) compute the extensions; and (2) compute the dispute tree [22] needed to prove it. The dispute tree is depicted in Figure 4 which suggests that only the argument network comprising **c**, **d**, and **b** needs to be considered.

#### 4.4 Explaining Extensions

Explaining multiple extensions is analogous to present different argumentation (sub-)networks that are linked together by attacks. Therefore, the solution for this communication goal builds upon the procedures envisaged in Sections 4.2 and 4.3.

It is worth mentioning that a possible way to identify the attack connections between those extensions is to consider a set of arguments belonging to a set of issues  $S \in \mathcal{E}_{\text{issues}}(\Gamma)$  such that  $|S| = |\mathcal{E}_\sigma(\Gamma)|$  for the chosen  $\sigma$  semantics.<sup>7</sup> Those arguments could provide foci of attention: for instance,  $\{\mathbf{b}, \mathbf{d}\} \in \mathcal{E}_{\text{issues}}(\Gamma_m)$ . Therefore, the text presented in Section 4.2 could be adapted in a straightforward manner to communicate the existence of two preferred extensions that gravitate around the issue  $\{\mathbf{b}, \mathbf{d}\}$ .

## 5 Conclusion

In this paper we have provided a blueprint for a complex set of research questions that arise when considering how to generate natural language representations of formal argumentation structures.

First of all, as already noted at the beginning of Section 4, we assume neither a model of the user, nor pre-existing contexts that need to be referenced in the piece of text generated. These elements will need to be addressed in future research.

Moreover some pieces of generated text sound quite awkward. To address such an issue, each piece of information inserted in an argument network should represent a single *text-agnostic normal form* proposition. This is clearly a constraint that might be unnatural if an untrained user tries to generate an argument network, and this raises an interesting research question concerning how to enable untrained users to formalise their lines of reasoning.

It is worth noting that the research questions highlighted in this paper have been grounded on a piece of argumentation formalised from a discussion between humans. We claim neither that our list is exhaustive nor that it is applicable to all possible argument networks. For instance, providing a natural language interface to an argument network built by an expert might lead to different, unforeseen, communicative goals. For instance, related research enabling the scrutiny of autonomous systems by allowing agents to generate plans through argument and dialogue [30] had the specific communication goal of justifying the purpose of each step within a plan.

Moreover we support the effort of extending AIF to include additional information: for instance, the AIFdb [20] project extends the AIF model to include schemes: of rephrasing; of locution describing communicative intentions which speakers use to

<sup>7</sup>Assuming that the semantics can be related to the notion of a complete labelling (Def. 4).

introduce propositional contents; and of interaction or protocol describing relations between locutions. Dealing with those pieces of information raise further research questions. Indeed, as we saw in Section 3.3, arguments **a**, **b**, and **c** form a conflict-free set, and in Section 4 we often considered them altogether in an approximate argument. However, while argument **c** was put forward by Claire Fox in the original dialogue (Section 2), arguments **a** and **b** belong to Nick Dearden. The question here is how to include the sources of those arguments in the generated natural language text.

Finally, we will investigate the potential of applying NLG to existing systems using formal argumentation in real-world applications, such as CISpaces [31, 9, 8] — a system for collaborative intelligence analysis — and ArgMed [18, 35] — a system for reasoning about the results of clinical trials.

## References

1. Amgoud, L., Cayrol, C., Lagasquie-Schiex, M.C., Livet, P.: On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems* 23, 1062–1093 (2008)
2. Baroni, P., Caminada, M., Giacomin, M.: An introduction to argumentation semantics. *Knowledge Engineering Review* 26(4), 365–410 (2011)
3. Besnard, P., Hunter, A.: Constructing argument graphs with deductive arguments: a tutorial. *Argument & Computation* 5(1), 5–30 (2014)
4. Black, E., Hunter, A.: A Relevance-theoretic Framework for Constructing and Deconstructing Enthymemes. *Journal of Logic and Computation* 22(1), 55–78 (2012)
5. Booth, R., Caminada, M., Podlaszewski, M., Rahwan, I.: Quantifying Disagreement in Argument-based Reasoning. In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)* (2012)
6. Booth, R., Caminada, M.W., Dunne, P.E., Podlaszewski, M., Rahwan, I.: Complexity Properties of Critical Sets of Arguments. In: *Computational Models of Argument: Proceedings of COMMA 2014*. pp. 173–184 (2014)
7. Caminada, M.: On the Issue of Reinstatement in Argumentation. In: *Proceedings of the 10th European Conference in Artificial Intelligence (JELIA 2006)*. pp. 111–123 (2006)
8. Cerutti, F., Norman, T.J., Toniolo, A.: A tool to highlight weaknesses and strengthen cases: Cispaces.org. In: *Legal Knowledge and Information Systems - JURIX 2018: The Thirty-first Annual Conference, Groningen, The Netherlands, 12-14 December 2018*. pp. 186–189 (2018), <https://doi.org/10.3233/978-1-61499-935-5-186>
9. Cerutti, F., Norman, T.J., Toniolo, A., Middleton, S.E.: Cispaces.org: From fact extraction to report generation. In: *Computational Models of Argument - Proceedings of COMMA 2018, Warsaw, Poland, 12-14 September 2018*. pp. 269–280 (2018), <https://doi.org/10.3233/978-1-61499-906-5-269>
10. Cerutti, F., Tintarev, N., Oren, N.: Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In: *21st European Conference on Artificial Intelligence*. pp. 207–212 (2014)
11. Chesnevar, C.I., McGinnis, J., Modgil, S., Rahwan, I., Reed, C., Simari, G.R., South, M., Vreeswijk, G.A.W., Willmot, S.: Towards an argument interchange format. *The Knowledge Engineering Review* 21(04), 293 (2006)
12. Dung, P.M.: 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)
13. Dunne, P.E., Wooldridge, M.: Complexity of abstract argumentation. In: Rahwan, I., Simari, G. (eds.) *Argumentation in AI*, chap. 5, pp. 85–104. Springer-Verlag (2009)

14. Gabbay, D.: Fibring Argumentation Frames. *Studia Logica* 93(2/3), 231–295 (2009)
15. Gatt, A., Reiter, E.: SimpleNLG: A realisation engine for practical applications. In: *Proceedings of the 12th European Workshop on Natural Language Generation*. pp. 90–93. Association for Computational Linguistics (2009)
16. Grice, H.P.: Logic and Conversation. In: Cole, P., Morgan, J. (eds.) *Syntax and Semantics 3: Speech Acts*, pp. 41–57 (1975)
17. Grice, P.: *Studies in the Way of Words*, vol. 65 (1989)
18. Hunter, A., Williams, M.: Aggregating evidence about the positive and negative effects of treatments. *Artificial intelligence in medicine* 56(3), 173–90 (nov 2012)
19. Laronge, J.A.: Evaluating universal sufficiency of a single logical form for inference in court. *Law, Probability and Risk* 11(2-3), 159–196 (2012)
20. Lawrence, J., Bex, F., Reed, C., Snaith, M.: AIFdb: Infrastructure for the Argument Web. In: *COMMA*. pp. 515 – 516 (2012)
21. Mann, W.C., Thompson, S.A.: Rhetorical Structure Theory: Toward a functional theory of text organisation. *Text* 8, 243–281 (1988)
22. Modgil, S., Caminada, M.: Proof Theories and Algorithms for Abstract Argumentation Frameworks. In: Simari, G., Rahwan, I. (eds.) *Argumentation in Artificial Intelligence*, chap. Proof Theo, pp. 105–129. Springer US, Boston, MA (2009)
23. Paglieri, F., Castelfranchi, C.: Trust, relevance, and arguments. *Argument & Computation* 5(2-3), 216–236 (2014)
24. Rahwan, I., Reed, C.: The Argument Interchange Format. In: *Argumentation in Artificial Intelligence*, pp. 383–402. Springer US, Boston, MA (2009)
25. Reiter, E., Dale, R.: *Building Natural Language Generation Systems*. No. Cambridge (2006)
26. Rosenfeld, A., Kraus, S.: Providing arguments in discussions on the basis of the prediction of human argumentative behavior. *ACM Transactions on Interactive Intelligent Systems (TiIS)* 6(4), 30 (2016)
27. Rosenfeld, A., Kraus, S.: Strategical argumentative agent for human persuasion. In: *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI)*. vol. 285, p. 320. IOS Press (2016)
28. Sperber, D., Wilson, D.: *Relevance : communication and cognition* 2nd Ed. Blackwell Publishers (1995)
29. Sperber, D., Wilson, D.: *Relevance Theory*. In: *The Handbook of Pragmatics* (2005)
30. Tintarev, N., Kutlak, R., Oren, N., Deemter, K.V., Green, M., Masthoff, J., Vasconcelos, W.: SAsSy - Scrutable Autonomous Systems. In: *Do-Form: Enabling Domain Experts to use Formalised Reasoning* (2013)
31. Toniolo, A., Norman, T.J., Etuk, A., Cerutti, F., Ouyang, R.W., Srivastava, M., Oren, N., Dropps, T., Allen, J.A., Sullivan, P.: Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*. pp. 781—789 (2015)
32. van Eemeren, F and Grootendorst, R.: *A Systematic Theory of Argumentation*. No. 1 (2014)
33. Walton, D., Reed, C., Macagno, F.: *Argumentation schemes*. Cambridge University Press, NY (2008)
34. Walton, D.: *Argumentation Theory: A Very Short Introduction*. In: *Argumentation in Artificial Intelligence*, pp. 1–22. Springer US, Boston, MA (2009)
35. Williams, M., Liu, Z.W., Hunter, A., Macbeth, F.: An updated systematic review of lung chemo-radiotherapy using a new evidence aggregation method. *Lung cancer (Amsterdam, Netherlands)* 87(3), 290–5 (mar 2015)