

Automatic evaluation of design alternatives with quantitative argumentation

Pietro Baroni* , Marco Romano^a, Francesca Toni^b , Marco Aurisicchio^c  and Giorgio Bertanza^d 

^aDip. Ingegneria dell'Informazione, University of Brescia, Brescia, Italy; ^bDepartment of Computing, Imperial College London, UK; ^cDepartment of Mechanical Engineering, Imperial College London, London, UK; ^dDip. Ingegneria Civile, Architettura, Territorio, Ambiente e Matematica, University of Brescia, Brescia, Italy

(Received 15 May 2014; accepted 21 October 2014)

This paper presents a novel argumentation framework to support Issue-Based Information System style debates on design alternatives, by providing an automatic quantitative evaluation of the positions put forward. It also identifies several formal properties of the proposed quantitative argumentation framework and compares it with existing non-numerical abstract argumentation formalisms. Finally, the paper describes the integration of the proposed approach within the design Visual Understanding Environment software tool along with three case studies in engineering design. The case studies show the potential for a competitive advantage of the proposed approach with respect to state-of-the-art engineering design methods.

Keywords: argumentation; debate; design rationale; Issue-Based Information System (IBIS); decision support

Engineering design is often described as an information-processing activity based on problem-solving within the constraints of bounded rationality (Simon, 1996; Simon & Newell, 1971). It consists of decomposing an initial problem into a range of sub-problems, proposing and assessing partial solutions, and integrating them as to satisfy the overall problem. This process is collaborative and often involves communication between non-co-located engineers. The development and communication of design solutions require engineers to form and share design rationale, that is, the argumentation in favour or against proposed designs.

These aspects of the engineering design process have led to the development (Kunz & Rittel, 1970) and subsequent investigation (Buckingham Shum & Hammond, 1994; Fischer, Lemke, McCall, & Morch, 1991) of the *issue-based information system* (IBIS) method, a graph-based formalisation of the decisions made during a design process along with the reasons why they were made. The IBIS method envisions a decision-making process where problems (or issues) are given solutions (or answers) after a thorough debate involving technical, economical, life, environmental and safety considerations. It also provides means to actively develop, communicate and record the reasons (or arguments) in favour or against the options explored during the design process. Initially, IBIS was conceived purely as a conceptual information system and its first implementations were paper-based and totally operated by hand. However, over time several software tools supporting editing and visualisation of IBIS graphs have been developed, for example, Compendium, DRed and design Visual Understanding Environment (designVUE) (e.g. see Aurisicchio & Bracewell, 2013; Buckingham Shum et al., 2006). These IBIS-based tools,

*Corresponding author. Email: pietro.baroni@unibs.it

including designVUE, which was selected as a starting point for this research, still leave to the users the burden of actually deriving any conclusion from the argumentative process and, eventually, making a decision. This is a task that, depending on the structure of the graph, may not be trivial.

This paper describes the outcome of collaborative research, involving experts of *engineering design* and *argumentation theory*, undertaken to overcome the limitations of standard design tools in general, and designVUE in particular. The ultimate goal of this research is to support engineers by providing them with a visual tool to automatically evaluate alternative design solutions and suggest the most promising answers to a design issue, given the underlying graph structure developed during the design process.

Since one of the main features of argumentation theory is evaluating arguments' acceptability (e.g. as in Cayrol & Lagasque-Schiex, 2005a; Dung, 1995) or strength (e.g. as in Cayrol & Lagasque-Schiex, 2005b; Evripidou & Toni, 2012; Leite & Martins, 2011; Matt & Toni, 2008) within debates and dialogues, we have singled it out as a promising companion to engineering design to achieve our research goal. For this application area, conventional notions of 'binary' acceptability (e.g. the notions in Dung, 1995), sanctioning arguments as acceptable or not, are better replaced with notions of numerical strength, as the latter are more fine-grained and allow to distinguish different degrees of acceptability.

This paper presents both theoretical and practical results. On the theoretical side, we propose a formal method to assign a numerical score to the nodes of an IBIS graph, starting from a base score provided by users. On the practical side, we describe the implementation of this method within designVUE and its preliminary evaluation in the context of three case studies.

The paper is organised as follows. Section 1 gives the basic notions concerning IBIS and the necessary background on argumentation theory. Section 2 introduces a form of argumentation frameworks abstracting away (a restricted form of) IBIS graphs, and Section 3 defines our approach for the quantitative evaluation of arguments in these frameworks. Section 4 studies some formal properties of our approach, and Section 5 gives formal comparisons with two traditional non-numerical argumentation frameworks, namely abstract (Dung, 1995) and bipolar (Cayrol & Lagasque-Schiex, 2005a) argumentation frameworks. Section 6 describes an implementation of our approach as an extension of designVUE, and Section 7 illustrates its application in three engineering case studies. Section 8 discusses related work, and Section 9 concludes.

The paper expands the work in Baroni, Romano, Toni, Aurisicchio, and Bertanza (2013) in several ways, notably by studying the properties of our proposed method for the quantitative evaluation of debates (Section 4), by considering the formal relationship with two traditional non-numerical argumentation frameworks (Section 5) and by developing one additional case study (Section 7.3). The latter amounts to revisiting a well-known design problem in the engineering design literature (Ulrich & Eppinger, 2004) and comparing it with standard decision techniques used for this problem, namely decision matrices (Pugh, 1991).

1. Background

1.1. Issue-Based Information System

IBIS (Kunz & Rittel, 1970) is a method to propose answers to issues and assess them through arguments. At the simplest level, the IBIS method consists of a structure that can be represented as a directed acyclic graph with four types of node: an *issue* node represents a problem being discussed, namely a question in need of an answer; an *answer* node represents a candidate solution to an issue; a *pro-argument* node represents an approval to a given answer or to another argument; a *con-argument* node represents an objection to a given answer or to another argument. An

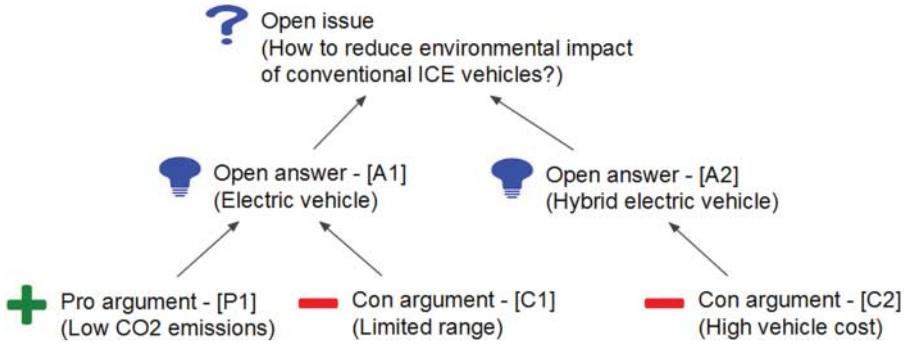


Figure 1. A simple IBIS graph.

answer node is always linked to an issue node, whereas pro-argument and con-argument nodes are normally linked to an answer node or to another argument. Each link is directed, pointing towards the dependent node.

Figure 1 shows an example of the IBIS graph, as implemented in designVUE, with a concrete illustration of the content of the nodes (labelled A1, A2, P1, C1 and C2, for convenience of reference) in the design domain of internal combustion engines (ICE). All the IBIS graphs presented hereafter are screenshots from the designVUE tool. This example graph has three layers: the first layer consists of an issue node, the second layer of two alternative answers and the third layer of arguments.

An IBIS graph is typically constructed according to the following rules: (1) an issue is captured; (2) answers are laid out and linked to the issue; (3) arguments are laid out and linked to either the answers or other arguments and (4) further issues may emerge during the process and be linked to either the answers or the arguments.

Conceptually, the addition of an answer or an argument corresponds to a move in the exploration of the design space.

In the class of design problems we considered for our application, IBIS graphs have specific features. First, each graph concerns a single issue (but this may involve addressing several sub-issues in turn). Second, answers correspond to alternative, mutually incompatible, solutions which can satisfy or not the dependent issue. Each answer is meant to represent a full solution to the issue hence they are mutually incompatible. Typically multiple satisfactory solutions are possible and can be accepted. Argumentation is used to screen them and select just one solution to be put forward. This differs from applications in other domains, for example, in diagnosis, where a combination of different answers may provide the cause for a fault.

In the designVUE implementation of the IBIS method (Aurisicchio & Bracewell, 2013), the four nodes can have alternative statuses to help users visualise aspects of the decision-making process (Figure 2).

The precise meaning of these statuses depends on the node type and is manually assigned by the users. For example, a designer may change the status of an answer from ‘open’ to ‘accepted’. In this paper, we define a method for automatic, rather than manual, evaluation of nodes in (restricted kinds of) IBIS graphs, based on argumentation theory, reviewed next.

1.2. Abstract argumentation and argument valuations

In this work, we will make use of abstract argumentation (Dung, 1995) and some extensions thereof. We review these briefly here.



Figure 2. Possible statuses of IBIS nodes.

Definition 1.1. A (finite) *abstract argumentation framework* (AF) is a pair $\langle \mathcal{X}, \mathcal{D} \rangle$, where \mathcal{X} is a finite set of *arguments* and $\mathcal{D} \subseteq \mathcal{X} \times \mathcal{X}$ is the *attack* (or *defeat*) relation. $(x, y) \in \mathcal{D}$ is referred to as ‘ x is an *attacker* (or *defeater*) of y ’.

An AF can be described as a directed graph whose nodes represent arguments and whose edges represent attacks. The nature and underlying structure of the arguments are completely abstracted away and the focus of the theory is essentially on the management of the conflicts represented by the attack relation. In this context, an (*argumentation*) *semantics* is a criterion to identify the *extensions* of an AF, namely those sets of arguments which can ‘survive the conflict together’. In turn, the *justification status* of an argument, according to a given semantics, can be defined in terms of its membership to the extensions prescribed by the semantics. A variety of semantics have been considered in the literature, whose review is beyond the scope of this paper (see Baroni, Caminada, & Giacomin, 2011 for a survey). These semantics evaluate arguments based on a binary notion of membership and thus give rise to a discrete set of justification statuses. We review here one of these semantics:

Definition 1.2. Given an AF $\mathcal{F} = \langle \mathcal{X}, \mathcal{D} \rangle$, a set $S \subseteq \mathcal{X}$ is *conflict-free* iff $\nexists a, b \in S : (a, b) \in \mathcal{D}$. A set S *defends* an argument a (or a is *acceptable* w.r.t. S) iff $\forall (b, a) \in \mathcal{D} \exists c \in S : (c, b) \in \mathcal{D}$. The *characteristic function* of \mathcal{F} is the mapping $F_{\mathcal{F}} : 2^{\mathcal{X}} \rightarrow 2^{\mathcal{X}}$ such that, for $S \subseteq \mathcal{X}$, $F_{\mathcal{F}}(S) = \{a \mid a \text{ is acceptable w.r.t. } S\}$. S is the *grounded extension* of \mathcal{F} , denoted as $GE(\mathcal{F})$, if S is the (unique) least fixed point of $F_{\mathcal{F}}$, that is, $S = F_{\mathcal{F}}(S)$ and there is no $S' \subsetneq S$ such that $S' = F_{\mathcal{F}}(S')$.

In the case of finite frameworks (as in the present paper), the grounded extension corresponds to the result of the iterative application of the characteristic function starting from the empty set until a fixed point is reached: $GE(\mathcal{F}) = \bigcup_{i \geq 1} F_{\mathcal{F}}^i(\emptyset)$, where, for any S , $F_{\mathcal{F}}^1(S) = F_{\mathcal{F}}(S)$ and $F_{\mathcal{F}}^i(S) = F_{\mathcal{F}}(F_{\mathcal{F}}^{i-1}(S))$ for $i > 1$.

While AFS are focused on conflicts between arguments, other forms of argument interaction can be considered, in particular a support relation, which can be incorporated into AFS to give rise to bipolar AFS (Cayrol & Lagasque-Schiex, 2005a):

Definition 1.3. A (finite) *bipolar AF* (BAF) is a triple $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$, where $\langle \mathcal{X}, \mathcal{D} \rangle$ is a (finite) AF and $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{X}$ is the *support relation*. A pair $(x, y) \in \mathcal{S}$ is referred to as ‘ x is a *supporter* of y ’.

The discrete argument evaluation for AFS can be extended to BAFs (see Cayrol & Lagasque-Schiex, 2005a). We recall the notion of stable extension for BAFs given in Section 6.2 of Amgoud, Cayrol, Lagasque-Schiex, & Livet (2008).

Definition 1.4. Given a BAF $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$ with $a, b \in \mathcal{X}$ a path of length $n > 1$, from a to b is a sequence a_1, \dots, a_n with $a_1 = a$, $a_n = b$ and $\forall 1 \leq i < n$ $(a_i, a_{i+1}) \in \mathcal{D} \vee (a_i, a_{i+1}) \in \mathcal{S}$. There is a *supported defeat* from an argument a to an argument b if there is a path with length $n \geq 3$ from a to b such that $\forall 1 \leq i < n - 1$ $(a_i, a_{i+1}) \in \mathcal{S}$ and $(a_{n-1}, a_n) \in \mathcal{D}$. A set $S \subseteq \mathcal{X}$ *defeats* an argument b if $\exists a \in S$ such that $(a, b) \in \mathcal{D}$ or there is a supported defeat from a to b . A set S is *conflict-free* iff $\nexists a, b \in S$ such that $\{a\}$ defeats b . A set S is a *stable extension* if S is conflict-free and $\forall a \in (\mathcal{X} \setminus S)$ S defeats a .

Another direction of enhancement of AFS amounts to assigning a numerical evaluation to arguments on a continuous scale. We recall here two proposals in this direction. The first gives a notion of local gradual valuation of a BAF that can be summarised as follows (see Cayrol & Lagasque-Schiex, 2005b for details):

Definition 1.5. Let L be a completely ordered set, L^* be the set of all the finite sequences of elements of L (including the empty sequence $()$), and H_{def} and H_{sup} be two ordered sets. Let $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$ be a BAF. Then, a *local gradual valuation* on $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$ is a function $v : \mathcal{X} \rightarrow L$ such that, for a generic argument $a \in \mathcal{X}$, given $\mathcal{D}^-(a) = \{d_1, \dots, d_n\}$, the set of attackers of a , and $\mathcal{S}^-(a) = \{s_1, \dots, s_p\}$, the set of supporters of a (for $n, p \geq 0$):

$$v(a) = g(h_{sup}(v(s_1), \dots, v(s_p)), h_{def}(v(d_1), \dots, v(d_n))),$$

where $g : H_{sup} \times H_{def} \rightarrow L$ is a function with $g(x, y)$ increasing on x and decreasing on y , and $h_{def} : L^* \rightarrow H_{def}/h_{sup} : L^* \rightarrow H_{sup}$ are functions (valuing the quality of the defeat/support, respectively) satisfying for any x_1, \dots, x_m, x_{m+1} (here $h = h_{def}$ or h_{sup}): (i) if $x_i \geq x_i'$ then $h((x_1, \dots, x_i, \dots, x_m)) \geq h((x_1, \dots, x_i', \dots, x_m))$; (ii) $h((x_1, \dots, x_m)) \leq h((x_1, \dots, x_m, x_{m+1}))$; (iii) $h() \leq h((x_1, \dots, x_n))$ and (iv) $h((x_1, \dots, x_n))$ is bounded by a limit value β .

Note that the local gradual valuation (LGV in the remainder) of an argument is defined recursively in terms of the valuations of its attackers and supporters.

The second proposal we consider is the extended social abstract argumentation approach of Evripidou & Toni (2012), taking into account, in addition to attackers and supporters, also positive and/or negative votes on arguments. In a nutshell, the idea is that in a social context (like an Internet-based social network or debate) opinions (arguments) are evaluated by a community of users through a voting process.

Definition 1.6. An *Extended Social Abstract Argumentation Framework* (ESAAF) is a 4-tuple $\langle \mathcal{X}, \mathcal{D}, \mathcal{S}, \mathcal{V} \rangle$, where $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$ is a (finite) BAF and $\mathcal{V} : \mathcal{X} \rightarrow \mathbb{N} \times \mathbb{N}$ is a function mapping arguments to the number of their positive and negative votes.

Given an (acyclic) ESAAF, argument evaluation is based on votes and on the attack/support relations. It involves a set of operators extending those of Leite & Martins (2011), where only attackers were considered. Omitting details, informally, this approach is based on defining a *semantic framework* in terms of a number of operators, some of which are quickly recalled below for the sake of comparison with LGV. In particular, the operator τ evaluates the social support for each argument a , based on its accumulated positive and negative votes (given by \mathcal{V}), and so assigns an *initial score*, $\tau(a)$, to a . This initial score has no counterpart in LGV seen earlier. Then, as in the case of LGV, the valuation of a is defined recursively in terms of the valuations of its attackers and supporters. The individual valuations of the attackers and of the supporters of a are first aggregated using the \vee operator. Then the aggregated valuations of the attackers and supporters are combined with $\tau(a)$. This results in a pair of values which roughly corresponds to the pair $h_{sup}((v(s_1), \dots, v(s_p))), h_{def}((v(d_1), \dots, v(d_n)))$ in LGV, the main difference being the

fact that $\tau(a)$ can be regarded as an additional parameter of these functions. Finally, the \uplus operator maps the above pair of values in a single final evaluation (and so clearly corresponds to the function g in LGV).

2. Quantitative argumentation debate frameworks

In Section 1.1, we have seen that the design scenarios we consider require IBIS graphs with specific features, and in particular with a single specific (design) issue and answers (linking to that issue) corresponding to different alternative solutions. Whereas IBIS graphs (in general and in design contexts) allow new issues to be brought up during the argumentation, as sub-issues of the main issue that are being debated, in this paper for simplicity we will disallow this possibility and focus on design debates that can be represented by IBIS graphs where arguments can only be pointed to by other arguments. Moreover, we focus on graphs in the restricted form of trees, with issues as roots.

We will define, in Section 3, a method for evaluating arguments and answers in IBIS graphs of the restricted kind we consider, aimed at accompanying or replacing the manual evaluation available in some IBIS implementations (Section 1.1). Examining some design scenarios with the relevant experts (see also Section 7), it emerged that, in their valuations, they typically ascribe different importance to arguments, which entails that a *base score* is required as a starting point for the evaluation. To fulfil these requirements, we propose a formal framework as follows:

Definition 2.1. A *QuAD (quantitative argumentation debate) framework* is a 5-tuple $(\mathcal{A}, \mathcal{C}, \mathcal{P}, \mathcal{R}, BS)$ such that (for *scale* $\mathbb{I}=[0, 1]$) :

- \mathcal{A} is a finite set of *answer arguments*;
- \mathcal{C} is a finite set of *con-arguments*;
- \mathcal{P} is a finite set of *pro-arguments*;
- the sets \mathcal{A} , \mathcal{C} , and \mathcal{P} are pairwise disjoint;
- $\mathcal{R} \subseteq (\mathcal{C} \cup \mathcal{P}) \times (\mathcal{A} \cup \mathcal{C} \cup \mathcal{P})$ is an acyclic binary relation;
- $BS : (\mathcal{A} \cup \mathcal{C} \cup \mathcal{P}) \rightarrow \mathbb{I}$ is a total function; $BS(a)$ is the *base score* of a .

The framework is referred to as ‘quantitative’ due to the presence of the base score. Ignoring this score, clearly QuAD frameworks are abstractions of (restricted forms of) IBIS graphs, with the issue node omitted since QuAD frameworks focus on the evaluation of answer nodes for a specific (implicit) issue. For example, the QuAD representation of the IBIS graph in Figure 1 has $\mathcal{A} = \{A1, A2\}$, $\mathcal{C} = \{C1, C2\}$, $\mathcal{P} = \{P1\}$ and $\mathcal{R} = \{(P1, A1), (C1, A1), (C2, A2)\}$. Note that QuAD frameworks may always be represented as sets of trees (one for each answer) . For example, the QuAD framework with $\mathcal{A} = \{A1, A2\}$, $\mathcal{C} = \{C1\}$, $\mathcal{P} = \{P1\}$ and $\mathcal{R} = \{(P1, A1), (P1, A2), (C1, A2), (C1, P1)\}$ corresponds to the two trees in Figure 3.

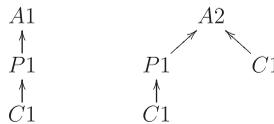


Figure 3. Tree representation of an example QuAD framework.

It is easy to see that a QuAD framework can also be interpreted as a BAF (again ignoring the base score), as notions of attack and support are embedded in the disjoint sets \mathcal{C} and \mathcal{P} . This is made explicit by the following definition.

Definition 2.2. Let $\mathcal{F} = \langle \mathcal{A}, \mathcal{C}, \mathcal{P}, \mathcal{R}, \mathcal{BS} \rangle$ be a QuAD framework and let $a \in \mathcal{A} \cup \mathcal{C} \cup \mathcal{P}$. The set of *direct attackers of a* is defined as $\mathcal{R}^-(a) = \{b \in \mathcal{C} | (b, a) \in \mathcal{R}\}$. The set of *direct supporters of a* is defined as $\mathcal{R}^+(a) = \{b \in \mathcal{P} | (b, a) \in \mathcal{R}\}$. Then, the BAF corresponding to \mathcal{F} is $\langle \mathcal{X}, \mathcal{D}, \mathcal{S} \rangle$ such that: $\mathcal{X} = \mathcal{A} \cup \mathcal{C} \cup \mathcal{P}$, $\mathcal{D} = \{(b, a) | b \in \mathcal{R}^-(a), a \in \mathcal{X}\}$, $\mathcal{S} = \{(b, a) | b \in \mathcal{R}^+(a), a \in \mathcal{X}\}$.

Note that an ESAAF equipped with a semantic framework can give rise to a QuAD framework, with the base score in the QuAD framework given by the initial score τ in the semantic framework for the ESAAF. The semantic framework includes, however, a recipe for calculating the initial score of arguments, based on votes in the ESAAF, whereas our QuAD framework assumes that the base score is given. Indeed, differently from the application contexts envisaged for ESAAF, design debates do not involve large communities of users so the notion of a base score based on votes does not seem to be appropriate, rather the base score can be represented as a numerical value that, for example, can be directly assessed by experts, or derived from information on the importance of criteria that arguments assess.

As we will see in Section 7, the choice of base scores for arguments is important for a correct evaluation outcome and far from simple since it has to take into account some case-specific factors: the definition of a methodology for assessing these scores based on application features is an important direction for future work.

3. Automatic evaluation in QuAD frameworks

Given a QuAD framework, in order to support the decision-making process by design engineers, we need a method to assign a quantitative evaluation, called *final score*, to answer nodes. To this purpose, we investigate the definition of a *score function* \mathcal{SF} for arguments. The basic idea is that the final score of an argument depends on its base score and on the final scores of its attackers and supporters, so \mathcal{SF} is defined recursively using a *score operator* combining these three elements.

We have defined direct attackers and supporters as *sets* (see Definition 2.2), taken from a (static) QuAD framework. However, in a dynamic design context these may actually be given in sequence. We will thus define the final score of an argument in terms of *sequences* of direct attackers and supporters. In this paper, we assume that these sequences are arbitrary permutations of the attackers and supporters (however, in a dynamic setting they may actually be given from the onset). For a generic argument a , let (a_1, \dots, a_n) be an arbitrary permutation of the $(n \geq 0)$ attackers in $\mathcal{R}^-(a)$. We denote as $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) = (\mathcal{SF}(a_1), \dots, \mathcal{SF}(a_n))$ the corresponding sequence of final scores. If $\mathcal{R}^-(a) = \{\}$, $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) = ()$, where $()$ denotes the empty sequence. Similarly, letting (b_1, \dots, b_m) be an arbitrary permutation of the $(m \geq 0)$ supporters in $\mathcal{R}^+(a)$, we denote as $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) = (\mathcal{SF}(b_1), \dots, \mathcal{SF}(b_m))$ the corresponding sequence of final scores, and, if $\mathcal{R}^+(a) = \{\}$, $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) = ()$. Finally, with an abuse of notation, $\mathcal{R}^-(a)$ and $\mathcal{R}^+(a)$ will stand also for their arbitrary permutations (a_1, \dots, a_n) and (b_1, \dots, b_m) , respectively, throughout the paper.

Using the hypothesis (implicitly adopted in Cayrol & Lagasque-Schiex, 2005b; Evripidou & Toni, 2012) of separability of the evaluations of attackers and supporters,¹ a generic score function for an argument a can be given as

$$\mathcal{SF}(a) = g(\mathcal{BS}(a), \mathcal{F}_{att}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))), \mathcal{F}_{supp}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))), \quad (1)$$

where g stands for (the second letter in) ‘aggregation’. Before we define g , \mathcal{F}_{att} and \mathcal{F}_{supp} , we illustrate the application of Equation (1) to the example of Figure 1. Suppose that $\mathcal{BS}(A1) = \mathcal{BS}(A2) = 0.5$, $\mathcal{BS}(C1) = 0.7$, $\mathcal{BS}(C2) = 0.4$, $\mathcal{BS}(P1) = 0.9$. Then, we obtain

$$\begin{aligned} \mathcal{SF}(A1) &= g(0.5, \mathcal{F}_{att}(0.5, SEQ_{\mathcal{SF}}((C1))), \mathcal{F}_{supp}(0.5, SEQ_{\mathcal{SF}}((P1)))); \\ \mathcal{SF}(A2) &= g(0.5, \mathcal{F}_{att}(0.5, SEQ_{\mathcal{SF}}((C2))), \mathcal{F}_{supp}(0.5, ()); \\ \mathcal{SF}(C1) &= g(0.7, \mathcal{F}_{att}(0.7, ()), \mathcal{F}_{supp}(0.7, ()); \\ \mathcal{SF}(C2) &= g(0.4, \mathcal{F}_{att}(0.4, ()), \mathcal{F}_{supp}(0.4, ()); \\ \mathcal{SF}(P1) &= g(0.9, \mathcal{F}_{att}(0.9, ()), \mathcal{F}_{supp}(0.9, ()). \end{aligned}$$

We identify some basic requirements for the score function \mathcal{SF} . First, each attacker (supporter) should have a negative or null (positive or null, respectively) effect on the final scores. Given any sequence $S = (s_1, \dots, s_k) \in \mathbb{I}^* \triangleq \bigcup_{i \geq 0} \mathbb{I}^i$ and $v \in \mathbb{I}$, let $S \cup (v)$ denote the sequence $(s_1, \dots, s_k, v) \in \mathbb{I}^{k+1}$. The above requirements can then be expressed, for sequences $S_1, S_2 \in \mathbb{I}^*$ and scores $v, v_0 \in \mathbb{I}$, as

$$g(v_0, \mathcal{F}_{att}(v_0, S_1), \mathcal{F}_{supp}(v_0, S_2)) \geq g(v_0, \mathcal{F}_{att}(v_0, S_1 \cup (v)), \mathcal{F}_{supp}(v_0, S_2)), \quad (2)$$

$$g(v_0, \mathcal{F}_{att}(v_0, S_1), \mathcal{F}_{supp}(v_0, S_2)) \leq g(v_0, \mathcal{F}_{att}(v_0, S_1), \mathcal{F}_{supp}(v_0, S_2 \cup (v))). \quad (3)$$

Further, we need to deal appropriately with those sequences which are *ineffective*, where a sequence Z is ineffective if it is empty or consists of all zeros. Formally the set of ineffective sequences is defined as $\mathcal{Z} = \bigcup_{i \geq 0} \{0\}^i$. Intuitively, when both the sequences of final scores of attackers and supporters are ineffective, the base score should remain unchanged. Formally, for every $Z_1, Z_2 \in \mathcal{Z}$, we require

$$g(v_0, \mathcal{F}_{att}(v_0, Z_1), \mathcal{F}_{supp}(v_0, Z_2)) = v_0. \quad (4)$$

In our running example, since $C1$, $C2$ and $P1$ have no attackers or supporters, we thus get $\mathcal{SF}(C1) = 0.7$; $\mathcal{SF}(C2) = 0.4$; $\mathcal{SF}(P1) = 0.9$.

To properly deal with ineffective sequences, we use a *special value nil* $\notin \mathbb{I}$ returned by \mathcal{F}_{att} and \mathcal{F}_{supp} . Formally, for every $v_0 \in \mathbb{I}$ and every sequence $Z \in \mathcal{Z}$, we impose

$$\mathcal{F}_{att}(v_0, Z) = \mathcal{F}_{supp}(v_0, Z) = nil. \quad (5)$$

For non-ineffective sequences, we define \mathcal{F}_{att} (and dually \mathcal{F}_{supp}) so that the contribution of an attacker (supporter) to the score of an argument decreases (increases) the argument score by an amount proportional both to (i) the score of the attacker (supporter), that is, a strong attacker (supporter) has more effect than a weaker one, and to (ii) the previous score of the argument itself, that is, an already strong argument benefits quantitatively less from a support than a weak one and an already weak argument suffers quantitatively less from an attack than a stronger one. These choices for \mathcal{F}_{att} and \mathcal{F}_{supp} are inspired by characteristics of human debate as well as the case studies we will present later and intuitively correspond to a sort of saturation of the effect of multiple attackers and supporters. Focusing on the case of a single attacker (supporter) with score $v \neq 0$, this leads to the following base expressions:²

$$f_{att}(v_0, v) = v_0 - v_0 \cdot v = v_0 \cdot (1 - v), \quad (6)$$

$$f_{supp}(v_0, v) = v_0 + (1 - v_0) \cdot v = v_0 + v - v_0 \cdot v. \quad (7)$$

The definitions of \mathcal{F}_{att} and \mathcal{F}_{supp} have then the same recursive form. Let $*$ stand for either *att* or *supp*. Then, for a non-ineffective sequence $S \in \mathbb{I}^*$:

$$\text{if } S = (v) : \mathcal{F}_*(v_0, S) = f_*(v_0, v) \quad (8)$$

$$\text{if } S = (v_1, \dots, v_n) : \mathcal{F}_*(v_0, (v_1, \dots, v_n)) = f_*(\mathcal{F}_*(v_0, (v_1, \dots, v_{n-1})), v_n). \quad (9)$$

Note that this definition directly entails that, for non-ineffective sequences S , $\mathcal{F}_{att}(v_0, S) \geq \mathcal{F}_{att}(v_0, S \cup (v))$ and $\mathcal{F}_{supp}(v_0, S) \leq \mathcal{F}_{supp}(v_0, S \cup (v))$. In our running example, we get

$$\mathcal{F}_{att}(0.5, SEQ_{\mathcal{SF}}((C1))) = \mathcal{F}_{att}(0.5, (0.7)) = f_{att}(0.5, 0.7) = 0.15;$$

$$\mathcal{F}_{supp}(0.5, SEQ_{\mathcal{SF}}((P1))) = \mathcal{F}_{supp}(0.5, (0.9)) = f_{supp}(0.5, 0.9) = 0.95;$$

$$\mathcal{F}_{att}(0.5, SEQ_{\mathcal{SF}}((C2))) = \mathcal{F}_{att}(0.5, (0.4)) = f_{att}(0.5, 0.4) = 0.3.$$

We now establish some basic properties of \mathcal{F}_{att} and \mathcal{F}_{supp} . First, unless the sequence is ineffective, they return values in $\mathbb{I} = [0, 1]$, as required:

Proposition 3.1. For any $v_0 \in \mathbb{I}$ and for any sequence $(v_1, \dots, v_k) \in \mathbb{I}^* \setminus \mathcal{Z}$, $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) \in \mathbb{I}$ and $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_k)) \in \mathbb{I}$.

Proof By induction on k . For the base case, trivially the statement holds for $k = 1$ given the definitions of f_{att} and f_{supp} . Assume that the statement holds for a generic sequence of length $k - 1$, that is, $\mathcal{F}_{att}(v_0, (v_1, \dots, v_{k-1})) = v_x \in \mathbb{I}$ then, from Equation (9), $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) = f_{att}(v_x, v_k)$. Similarly, letting $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_{k-1})) = v_y \in \mathbb{I}$, we get $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_k)) = f_{supp}(v_y, v_k)$. Then, again the statement holds by definition of f_{att} and f_{supp} . ■

Then, it is of course required that \mathcal{F}_{att} and \mathcal{F}_{supp} produce the same result for any permutation of the same sequence.

Proposition 3.2. For any $v_0 \in \mathbb{I}$ and $(v_1, \dots, v_k) \in \mathbb{I}^*$, let $(v_{1_i}, \dots, v_{k_i})$ be an arbitrary permutation of (v_1, \dots, v_k) . It holds that $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) = \mathcal{F}_{att}(v_0, (v_{1_i}, \dots, v_{k_i}))$ and $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_k)) = \mathcal{F}_{supp}(v_0, (v_{1_i}, \dots, v_{k_i}))$.

Proof Obvious for ineffective sequences. Otherwise, as to \mathcal{F}_{att} we note that $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) = f_{att}(f_{att}(\dots f_{att}(v_0, v_1) \dots), v_{k-1}, v_k) = (((v_0 \cdot (1 - v_1)) \cdot (1 - v_2)) \dots (1 - v_k)) = v_0 \cdot \prod_{i=1}^k (1 - v_i)$. Thus, the statement follows directly from commutativity and associativity of the product of the $(1 - v_i)$ factors. As to \mathcal{F}_{supp} , $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_k)) = f_{supp}(f_{supp}(\dots f_{supp}(v_0, v_1) \dots), v_{k-1}, v_k)$. Thus, the statement follows from the well-known properties of commutativity and associativity of any T-conorm. ■

Another desirable property of \mathcal{F}_{att} and \mathcal{F}_{supp} is a sort of monotonic behaviour with respect to the increasing score of attackers and supporters, respectively.

Proposition 3.3. For any $v_0 \in \mathbb{I}$ and for any $S = (v_1, \dots, v_h, \dots, v_k) \in \mathbb{I}^* \setminus \mathcal{Z}$, $1 \leq h \leq k$, let S^+ be a sequence obtained from S by replacing v_h with some $v_l > v_h$. Then $\mathcal{F}_{att}(v_0, S) \geq \mathcal{F}_{att}(v_0, S^+)$ and $\mathcal{F}_{supp}(v_0, S) \leq \mathcal{F}_{supp}(v_0, S^+)$.

Proof As to \mathcal{F}_{att} given that for a generic sequence $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) = v_0 \cdot \prod_{i=1}^k (1 - v_i)$, we observe that $\mathcal{F}_{att}(v_0, S^+) = \mathcal{F}_{att}(v_0, S) \cdot (1 - v_l)/(1 - v_h)$ and the statement follows from $0 \leq 1 - v_l < 1 - v_h$. As to \mathcal{F}_{supp} , from commutativity and associativity of f_{supp} , letting $S^* = (v_1, \dots, v_{h-1}, v_{h+1}, \dots, v_k) \in \mathbb{I}^{k-1}$, we get $\mathcal{F}_{supp}(v_0, S) = f_{supp}(\mathcal{F}_{supp}(v_0, S^*), v_h)$ and

$\mathcal{F}_{supp}(v_0, S^+) = f_{supp}(\mathcal{F}_{supp}(v_0, S^*), v_l)$ and the statement follows from the well-known monotonicity of T-conorms. ■

In order to finalise the definition of score function, we need to define g . For this we adopted the idea that when the effect of attackers is null (i.e. the value returned by \mathcal{F}_{att} is *nil*), the final score must coincide with the one established on the basis of supporters, and dually when the effect of supporters is null, while, when both are null, condition (4) applies. When both attackers and supporters have an effect, the final score is obtained averaging the two contributions. This amounts to treating the aggregated effect of attackers and supporters equally in determining the strength of the argument (see also the discussion in Section 4.2). Formally:

Definition 3.4. The operator $g : \mathbb{I} \times \mathbb{I} \cup \{nil\} \times \mathbb{I} \cup \{nil\} \rightarrow \mathbb{I}$ is defined as follows:

$$g(v_0, v_a, v_s) = v_a \text{ if } v_s = nil \text{ and } v_a \neq nil, \quad (10)$$

$$g(v_0, v_a, v_s) = v_s \text{ if } v_a = nil \text{ and } v_s \neq nil, \quad (11)$$

$$g(v_0, v_a, v_s) = v_0 \text{ if } v_a = v_s = nil, \quad (12)$$

$$g(v_0, v_a, v_s) = \frac{(v_a + v_s)}{2} \text{ otherwise.} \quad (13)$$

Then, the following result directly ensues from Propositions 3.1–3.3:

Proposition 3.5. The score function $\mathcal{SF}(a)$ defined by Equations (1), (8) and (9) and by Definition 3.4 satisfies properties (2)–(4).

For our running example, we get $\mathcal{SF}(A1) = g(0.5, 0.15, 0.95) = 0.55$ and $\mathcal{SF}(A2) = g(0.5, 0.3, nil) = 0.3$.

On the computational side, given that in a QuAD framework the relation \mathcal{R} is acyclic, evaluating \mathcal{SF} for answer nodes (in fact, for any node) is quite easy: given an argument a to be evaluated the score function is invoked recursively on its attackers and supporters to obtain $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))$ and $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))$ which are finally fed to the \mathcal{SF} operator along with the base score $\mathcal{BS}(a)$. The recursion is well-founded given the acyclicity of \mathcal{R} , the base case being provided by nodes with neither attackers nor supporters whose final score coincides with their base score.

4. Properties of the quantitative evaluation with \mathcal{SF}

In this section, we analyse some properties of the score function introduced in Section 3, namely the meaning of the extreme values 0 and 1 and the relevant behaviour of \mathcal{SF} , the range of possible values of the final score of an argument, and the result in the cases where the sets of attackers and supporters and their scores are symmetric.

4.1. Behaviour with extreme values

The extreme values 0 and 1 carry a specific meaning and should be used accordingly. Given an argument a , $\mathcal{BS}(a) = 0$ implies that the final evaluation of a is indifferent to attackers since $\mathcal{F}_{att}(0, S) = 0$ for every $S \in \mathbb{I}^* \setminus \mathcal{Z}$. Similarly $\mathcal{BS}(a) = 1$ implies that the final evaluation of a is indifferent to supporters since $\mathcal{F}_{supp}(1, S) = 1$ for every $S \in \mathbb{I}^* \setminus \mathcal{Z}$.

It can also be observed that an attacker with final score 1 has a *saturating* role as far as attackers against some argument a are concerned, since $\forall v_0 \in \mathbb{I}, \forall S \in \mathbb{I}^*, \mathcal{F}_{att}(v_0, S \cup \{1\}) = 0$, that is, the base score of a and any further attacker make no difference in this case. Similarly, a supporter with final score 1 has a saturating role, since $\forall v_0 \in \mathbb{I}, \forall S \in \mathbb{I}^*, \mathcal{F}_{supp}(v_0, S \cup \{1\}) = 1$.

It is also easy to see that either an attacker or a supporter with final score 0 has no effect on the final score and could be ignored. Indeed, $\forall v_0 \in \mathbb{I}, \forall S \in \mathbb{I}^*, \mathcal{F}_{att}(v_0, S \cup (0)) = \mathcal{F}_{att}(v_0, S)$ and $\mathcal{F}_{supp}(v_0, S \cup (0)) = \mathcal{F}_{supp}(v_0, S)$.

Further, extreme values cannot be attained in final scores unless some extreme values are present in the input values to \mathcal{SF} , as summarised by the following conditions, for any argument a :

$$\mathcal{BS}(a) \notin \{0, 1\} \wedge \nexists b \in \mathcal{R}^-(a) : \mathcal{SF}(b) \in \{0, 1\} \Rightarrow \mathcal{F}_{att}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))) \notin \{0, 1\}, \quad (14)$$

$$\mathcal{BS}(a) \notin \{0, 1\} \wedge \nexists b \in \mathcal{R}^+(a) : \mathcal{SF}(b) \in \{0, 1\} \Rightarrow \mathcal{F}_{supp}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))) \notin \{0, 1\}, \quad (15)$$

$$\begin{aligned} \mathcal{BS}(a) \notin \{0, 1\} \wedge \mathcal{F}_{att}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))) \notin \{0, 1\} \\ \wedge \mathcal{F}_{supp}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))) \notin \{0, 1\} \Rightarrow \mathcal{SF}(a) \notin \{0, 1\}. \end{aligned} \quad (16)$$

Given the acyclic structure of a QuAD-framework, it follows directly from the definition of \mathcal{SF} that extreme values can not be attained in final scores unless some extreme values are present in the base scores:

$$\nexists a \in \mathcal{X} : \mathcal{BS}(a) \in \{0, 1\} \Rightarrow \nexists a \in \mathcal{X} : \mathcal{SF}(a) \in \{0, 1\}. \quad (17)$$

The last property ensures in particular that extreme values (coherently with the special meaning they carry) may enter into play only by a deliberate choice of the expert providing the base scores.

4.2. Characterisation of the final score

In order to characterise the range of possible values of the final score, first note that the following inequalities hold, for any argument a assuming $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) \notin \mathcal{Z}$ and $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) \notin \mathcal{Z}$:

$$0 \leq \mathcal{F}_{att}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))) \leq \mathcal{BS}(a), \quad (18)$$

$$\mathcal{BS}(a) \leq \mathcal{F}_{supp}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^+(a))) \leq 1, \quad (19)$$

where the first inequality in Equation (18) and the second inequality in Equation (19) are strict in absence of extreme values as discussed previously, while the other inequalities, involving $\mathcal{BS}(a)$, are strict, respectively, in the cases where $\exists b \in \mathcal{R}^-(a) : \mathcal{SF}(b) > 0$ and $\exists b \in \mathcal{R}^+(a) : \mathcal{SF}(b) > 0$.

Applying Equation (13) from Definition 3.4 when $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) \notin \mathcal{Z}$ and $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) \notin \mathcal{Z}$, it follows that:

$$\frac{\mathcal{BS}(a)}{2} \leq \mathcal{SF}(a) \leq \frac{1 + \mathcal{BS}(a)}{2}. \quad (20)$$

The lower bound corresponds to an argument with very strong attackers and weak supporters: it has its base score halved. The upper bound corresponds to an argument with weak attackers and very strong supporters, for which the distance from one of the final score is half the one of the base score. This corresponds to the idea that differences in the base score assessments can only be reversed up to a certain extent as an effect of attackers and supporters. Note also that in case of a contradictory situation (both very strong attackers and very strong supporters), the final score is 0.5 independently of the base score.

Equation (13) applies in the ‘regular’ case where both the sets of attackers and supporters are non-empty and have some effect. The absence of any (effective) attacker or of any (effective)

supporter is treated as a special case in Definition 3.4 and this induces a discontinuity in the behaviour of the operator g (and hence of \mathcal{SF}): if an argument a has very strong attackers and no (effective) supporter at all, it may be the case that $\mathcal{SF}(a) = 0$, while adding even a single very weak supporter it turns out that $\mathcal{SF}(a) \geq \mathcal{BS}(a)/2$. A dual behaviour occurs for the case of attackers. This behaviour corresponds to the idea that the inability to indicate any (even weak) effective supporter (or attacker) for an argument is a peculiar situation justifying a drastic penalty (or reward) in the final evaluation. Whether this behaviour is suitable in all contexts is an open question, and the definition of different forms of \mathcal{SF} without this discontinuity is an important direction for future work.

4.3. Symmetry between attackers and supporters

One may wonder whether a symmetric configuration of attackers and supporters, that is, when the number of attackers and supporters is the same and they have pairwise equal strength, gives rise to a symmetric effect in the evaluation results. We show that a symmetry holds in \mathcal{F}_{att} and \mathcal{F}_{supp} concerning the distance from the extreme values. Intuitively, this means that if an argument a has a symmetric configuration of attackers and supporters \mathcal{F}_{att} reduces the distance of $\mathcal{BS}(a)$ from 0 in the same proportion as \mathcal{F}_{supp} reduces the distance of $\mathcal{BS}(a)$ from 1. This is shown by the following propositions.

Proposition 4.1. Let a be an argument with $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) = (v_1, \dots, v_k) \notin \mathcal{Z}$ for some $k \geq 1$. $(\mathcal{F}_{att}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a))) - 0) / (\mathcal{BS}(a) - 0) = \prod_{i=1}^k (1 - v_i)$.

Proof The proposition follows directly from $\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) = v_0 \cdot \prod_{i=1}^k (1 - v_i)$ (see the proof of Proposition 3.2). ■

Proposition 4.2. Let a be an argument with $SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) = (v_1, \dots, v_k) \notin \mathcal{Z}$ for some $k \geq 1$. $(1 - \mathcal{F}_{supp}(\mathcal{BS}(a), SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)))) / (1 - \mathcal{BS}(a)) = \prod_{i=1}^k (1 - v_i)$.

Proof The proof is by induction on the number k of supporters. For $k = 1$, observe that $\mathcal{F}_{supp}(v_0, (v_1)) = f_{supp}(v_0, v_1) = v_0 + v_1 - v_1 \cdot v_0 = 1 - (1 - v_1)(1 - v_0)$. Assuming inductively that for some $j \geq 1$, for every sequence $(v_1, \dots, v_j) \in \mathbb{I}^j$, $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_j)) = 1 - (1 - v_0) \cdot \prod_{i=1}^j (1 - v_i)$, we show that the same equality holds for every sequence $(v_1, \dots, v_{j+1}) \in \mathbb{I}^{j+1}$. In fact, $\mathcal{F}_{supp}(v_0, (v_1, \dots, v_{j+1})) = f_{supp}(\mathcal{F}_{supp}(v_0, (v_1, \dots, v_j)), v_{j+1}) = f_{supp}(1 - (1 - v_0) \cdot \prod_{i=1}^j (1 - v_i), v_{j+1}) = 1 - (1 - v_{j+1}) \cdot (1 - (1 - (1 - v_0) \cdot \prod_{i=1}^j (1 - v_i))) = 1 - (1 - v_{j+1}) \cdot (1 - v_0) \cdot \prod_{i=1}^j (1 - v_i) = 1 - (1 - v_0) \cdot \prod_{i=1}^{j+1} (1 - v_i)$. Then it follows $(1 - \mathcal{F}_{supp}(v_0, (v_1, \dots, v_{j+1}))) / (1 - v_0) = (1 - (1 - (1 - v_0) \cdot \prod_{i=1}^{j+1} (1 - v_i))) / (1 - v_0) = \prod_{i=1}^{j+1} (1 - v_i)$. ■

As a by-product of the proofs of the previous propositions, we observe that assuming $\mathcal{BS}(a) = v_0$ and $SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)) = SEQ_{\mathcal{SF}}(\mathcal{R}^+(a)) = (v_1, \dots, v_k)$, $k \geq 1$, we get $\mathcal{SF}(a) = (\mathcal{F}_{att}(v_0, (v_1, \dots, v_k)) + \mathcal{F}_{supp}(v_0, (v_1, \dots, v_k))) / 2 = (v_0 \cdot \prod_{i=1}^k (1 - v_i) + 1 - (1 - v_0) \cdot \prod_{i=1}^k (1 - v_i)) / 2 = (1 + (2v_0 - 1) \cdot \prod_{i=1}^k (1 - v_i)) / 2$. From this, it is easy to see that, in presence of symmetric attackers and supporters, $\mathcal{SF}(a)$ coincides with $\mathcal{BS}(a)$ only if $\mathcal{BS}(a) = 0.5$, that is, if the base score of a is in turn equidistant from 0 and 1. This shows that 0.5 is the correct base score value when there is no a priori attitude towards the acceptance or rejection of an argument.

5. Relationships with traditional non-numerical frameworks

In this section, we discuss the relationships of our approach with abstract argumentation frameworks (AFs) (Dung, 1995) and bipolar argumentation frameworks (BAFs) (Cayrol & Lagasque-Schiex, 2005a).

As to Dung's AFs, first note that since they encompass the relation of attack only and there are no arguments with a distinguished role, the set of arguments of Definition 1.1 corresponds to the set of con-arguments in Definition 2.1 and the attack relation of Definition 1.1 corresponds to the generic binary relation of Definition 2.1, under the constraint that it is acyclic. Further, acceptance is binary in AFs and there is no notion of base score. Accordingly, it can be observed that the implicit initial evaluation of arguments corresponds to full acceptance (it is well known that unattacked arguments are accepted in any Dung's semantics): this corresponds to assigning a base score of 1 to every argument. These considerations are summarised by the following definition.

Definition 5.1. Given an AF $\langle \mathcal{X}, \mathcal{D} \rangle$ such that \mathcal{D} is acyclic, the corresponding QuAD framework is defined as $\langle \emptyset, \mathcal{X}, \emptyset, \mathcal{D}, \mathcal{BS}_{\mathcal{X}}^1 \rangle$ where $\mathcal{BS}_{\mathcal{X}}^1 \triangleq \mathcal{X} \times \{1\}$.

As to the evaluation of arguments, in the case of an acyclic AF all traditional Dung's semantics prescribe exactly one extension (i.e. a set of justified arguments) coinciding with the *grounded* extension (Section 1.2). We prove that the QuAD framework corresponding to an acyclic AF \mathcal{F} and using the score function \mathcal{SF} assigns a final score of 1 to the members of $GE(\mathcal{F})$ and a final score of 0 to every other argument.

Proposition 5.2. Given an AF $\mathcal{F} = \langle \mathcal{X}, \mathcal{D} \rangle$ such that \mathcal{D} is acyclic and the corresponding QuAD framework $\langle \emptyset, \mathcal{X}, \emptyset, \mathcal{D}, \mathcal{BS}_{\mathcal{X}}^1 \rangle$, for every $a \in \mathcal{X} : \mathcal{SF}(a) = 1$ if $a \in GE(\mathcal{F})$, $\mathcal{SF}(a) = 0$ otherwise.

Proof First note that, since there are no supporters, for every argument a Equation (10) of Definition 3.4 applies, hence $\mathcal{SF}(a) = \mathcal{F}_{att}(\mathcal{BS}_{\mathcal{X}}^1(a), SEQ_{\mathcal{SF}}(\mathcal{R}^-(a)))$. Recalling that $GE(\mathcal{F}) = \bigcup_{i \geq 1} F_{\mathcal{F}}^i(\emptyset)$, we prove by induction that for every $i, \forall a \in F_{\mathcal{F}}^i(\emptyset), \mathcal{SF}(a) = 1$. As to the induction base, consider $F_{\mathcal{F}}^1(\emptyset) = F_{\mathcal{F}}(\emptyset)$: it consists of the arguments not receiving any attack. As such, from Equation (12), their final score is equal to the base score, namely 1. Suppose now that for every $i \geq 1, \forall a \in F_{\mathcal{F}}^i(\emptyset) \mathcal{SF}(a) = 1$ and consider $F_{\mathcal{F}}^{i+1}(\emptyset) = F_{\mathcal{F}}(F_{\mathcal{F}}^i(\emptyset))$. By the inductive hypothesis, $\forall a \in F_{\mathcal{F}}^{i+1}(\emptyset) \cap F_{\mathcal{F}}^i(\emptyset) \mathcal{SF}(a) = 1$. Considering any $a \in F_{\mathcal{F}}^{i+1}(\emptyset) \setminus F_{\mathcal{F}}^i(\emptyset)$ we have that a is defended by $F_{\mathcal{F}}^i(\emptyset)$, that is, $\forall b : (b, a) \in \mathcal{D} \exists c \in F_{\mathcal{F}}^i(\emptyset) : (c, b) \in \mathcal{D}$. By the inductive hypothesis, $\mathcal{SF}(c) = 1$ from which, using Equations (6) and (9), $\mathcal{F}_{att}(\mathcal{BS}_{\mathcal{X}}^1(b), SEQ_{\mathcal{SF}}(\mathcal{R}^-(b))) = 0$ follows, hence $\mathcal{SF}(b) = 0$. Using the same equations and taking into account that $\mathcal{BS}_{\mathcal{X}}^1(a) = 1$ we get $\mathcal{SF}(a) = 1$. To prove now that $\forall a \in \mathcal{X} \setminus GE(\mathcal{F}), \mathcal{SF}(a) = 0$, recall that it is well known that in acyclic AFs the grounded extension is also *stable*, that is, it attacks all other arguments. Then $\forall a \in \mathcal{X} \setminus GE(\mathcal{F}), \exists b \in GE(\mathcal{F}) : (b, a) \in \mathcal{D}$. Since, as just proved, $\mathcal{SF}(b) = 1$, using the same reasoning as above it follows $\mathcal{SF}(a) = 0$. ■

Turning to BAFs, one may wonder whether given a QuAD framework $\langle \mathcal{A}, \mathcal{C}, \mathcal{P}, \mathcal{R}, \mathcal{BS}_{\mathcal{AUCUP}}^1 \rangle$ the score function assigns a final score of 1 to the members of the unique stable extension (Section 1.2) of the corresponding BAF (Definition 2.2) and 0 to every other argument. The answer is negative due to the fact that the QuAD framework is based on the idea of a compensation between the effects of attackers and supporters, while in Definition 1.4 a defeat cannot be compensated by any support and support relations play basically only the role of 'defeat vehicles'.

For instance, considering the QuAD framework

$$\langle \{a\}, \{b, c\}, \{d\}, \{(b, a)\}, (c, b), (d, b), \mathcal{BS}_{\{a,b,c,d\}}^1 \rangle$$

and the corresponding BAF $\mathcal{B} = \langle \{a, b, c, d\}, \{(b, a), (c, b)\}, \{(d, b)\} \rangle$, we get that the stable extension of \mathcal{B} is $\{a, c, d\}$, but $\mathcal{SF}(a) = 0.5$, due to the fact that $\mathcal{SF}(b) = 0.5$ too. This is coherent with the logic-based notion of support adopted in BAFs, which is basically different from the one of pro-argument adopted in the QuAD framework. Recently, a variety of alternative interpretations of the binary notion of support in BAFs have been analysed in Cayrol & Lagasque-Schiex (2013), none of them, however, encompasses the compensation effect mentioned above, since, coherently with the logic-based perspective, in all cases an attack can only be countered by another attack.

6. Implementation in designVUE

The proposed approach has been implemented in designVUE, a pre-existing IBIS application.³ designVUE has been chosen as a platform for the implementation of the proposed approach for various reasons: it is open source; it has been developed by the Design Engineering Group at Imperial College London; it is receiving increasing interest from academia and industry and, as a result, has a growing user community. In the following paragraphs, we describe in more detail designVUE and its extension with the QuAD framework.

designVUE is an application developed using Java to attain cross-platform portability. Its GUI consists primarily of a main window, which contains the menu bar, the toolbar and the graph canvas.

The main purpose of designVUE is to draw graphs (also referred to as diagrams or maps) mostly consisting of nodes (depicted as boxes) and links (depicted as arrows) among them. The programme does not impose any restriction on the way a graph can be drawn. It is up to the user to confer any meaning to a graph. Among the large variety of graphs that can be drawn, designVUE supports IBIS graphs. These have no special treatment in designVUE and, in particular, there is no support to the evaluation of the argumentative process. In addition to the main window, there are floating windows that can be opened from the *Windows* menu. One of these, called *Info Window*, presents information about the currently selected node.

The QuAD framework has been implemented in Java and integrated into a customised version of designVUE, forking its existing codebase.⁴ The additions and modifications brought to designVUE fit broadly in two categories: those related to the GUI and those concerning the implementation of the score assignment method. As to the GUI:

- a new pane called *BaseScore Pane* has been added to the *Info Window*: it displays the base score of the currently selected IBIS node and allows the user to edit it (base scores are created with a default value of 0.5);
- a new pane called *Score Pane* has been added to *Info Window*: it displays the final score of the currently selected IBIS node;
- a new menu item labelled *Compute Argumentation on IBIS node* has been added to the *Content* menu: it can be invoked only after selecting an IBIS answer node and triggers the score computation for the selected node (and for all the nodes on which it depends).

Figure 4 shows a screenshot of the enhanced version of designVUE evidencing the above-mentioned features.

As to the algorithm to compute the final scores, it has been implemented in a Java class, which basically carries out a depth-first post-order traversal, which acts directly onto the IBIS nodes displayed in the canvas. To enhance performances in complex graphs where some pro- and/or con-arguments affect many other arguments (e.g. as in the example represented as trees

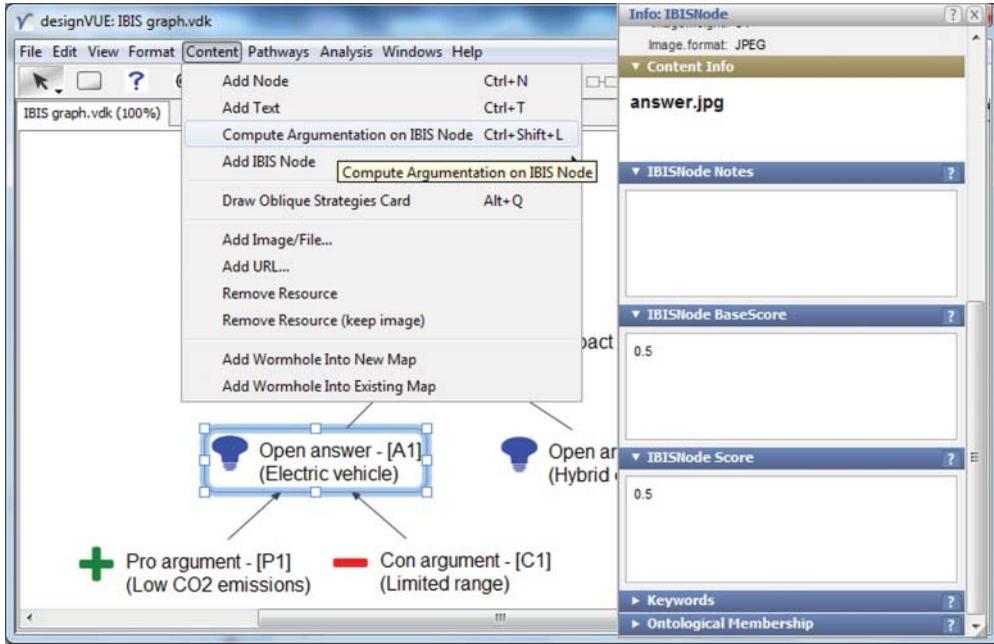


Figure 4. A screenshot of the enhanced version of designVUE.

in Figure 3), the algorithm implements a so-called closed list in order to reuse the scores already computed in previous phases of the graph traversal.

7. Case studies

A preliminary evaluation of the enhanced version of designVUE was carried out through three case studies. The first, in the domain of civil engineering, concerns the choice of a foundation for a multistorey building to be developed on a brownfield. This case study was developed in collaboration with a civil engineer with more than 10 years of experience in the industry, who was already familiar with the IBIS concept having used it through the Compendium software by Buckingham Shum et al. (2006). The second, in the domain of water engineering, focuses on the choice of a reuse technology for sludge produced by wastewater treatment plants. The third, in the domain of medical engineering, focuses on the design of an improved, reusable syringe with precise dosage control for outpatient use. This case study is a reformulation of a well-known design problem by Ulrich & Eppinger (2004).

The three cases are meant to explore the application of the proposed approach to decision problems with different structures: the foundations case features a canonical IBIS structure with rich rationale and a free-flowing debate, the sludge reuse case has a more rigid structure separated in two tiers where, similarly to the decision matrix approach, a fixed set of arguments is considered for each alternative, while the outpatient syringe case is based on a decision matrix example directly taken from the literature. Further, it can be noted that, differently from the others, the sludge reuse case concerns a decision process where both technical and non-technical considerations, drawn by different classes of actors, have to be taken into account.

The case studies also illustrate the use of the QuAD formalism in different application areas, where different conceptualisations and different nuances in the use of the base scores are adopted. They represent a preliminary investigation of the possible uses of the formalism, aimed at

collecting initial feedbacks from domain experts and to possibly point out major difficulties and drawbacks (which actually did not arise). While these cases were built incrementally by direct interaction between the developers and the domain experts, the development of a proper score elicitation and acquisition methodology is under way and represents a necessary prerequisite for an extensive validation.

7.1. Foundations

This case study is based on a design task, which was selected to satisfy the following criteria: the design problem had to be well known to the industry; and the problem-solving process had to rely on the application of known and established solution principles. On this basis the task presented in this case study can be considered to be at the boundary between adaptive and variant design (Pahl & Beitz, 1984). The reason for choosing this type of design task is to adopt a *walk before you run* approach to evaluation.

The case is based on real project experience of the collaborating engineer. However, it was not developed during the actual design process but rather reconstructed retrospectively. Prior to the development of the case, the engineer was introduced to the enhanced version of designVUE and instructed to use it including inputting values for the base scores.

As mentioned earlier, the design problem focuses on the selection of the most appropriate type of foundation for a multistorey building in a brownfield area. This is the part of urban planning concerning the reuse of abandoned or underused industrial and commercial facilities. When considering the choice of building foundations in brownfield sites, multiple alternatives are common and multiple considerations have to be made starting from the different kinds of ground and their load-bearing capabilities, which are usually different than in greenfield sites.

The starting point of the IBIS graph developed by the engineer is the issue to choose a suitable foundation given the requirements discussed earlier. Three types of foundation solutions are considered, namely Pad, Raft and Piles, and these are subsequently evaluated using several pro- and con-arguments (Figure 5).

After the development of the IBIS graph, the engineer executed the score computation on the three solutions under two situations: (1) using default values for the base scores and (2) using modified values for the base scores. The modified values for the base scores emerged through a three-step process involving extraction of the criteria behind each argument (see text in bracket at the bottom of each argument in Figure 5), analysis of the relative importance of the criteria in the context of the selected design task and assignment of a numerical value between 0 and 1 to each criterion matching the relative importance. As a result of this work the following base scores were assigned to the 10 criteria: performance and functional fulfilment (0.8); flexibility (0.4); additional structure (0.4); material use (0.3); buildability (0.2); cost (0.2); management complexity (0.1); execution complexity (0.1); unforeseen (0.1) and construction time (0.1).

The results for the situation with unchanged values indicate that Pad (0.51) is the preferred solution over Raft (0.49) and Piles (0.44). Differently, the results for the situation in which the values were changed suggest that Piles (0.56) is slightly preferable to Raft (0.55) and considerably preferable to Pad (0.41). As it can be seen, the three alternatives are ranked exactly in the reverse order. Only the results based on the modified values for the base scores were judged by the expert consistent with his conclusions.

On the one hand, this confirms the importance of weighting pro- and con-arguments with expert-provided base scores in order to get meaningful results. On the other hand, it shows that a purely graphical representation of the pros and cons is typically insufficient to give an account of the reasons underlying the final choice by the experts. In this sense, representing

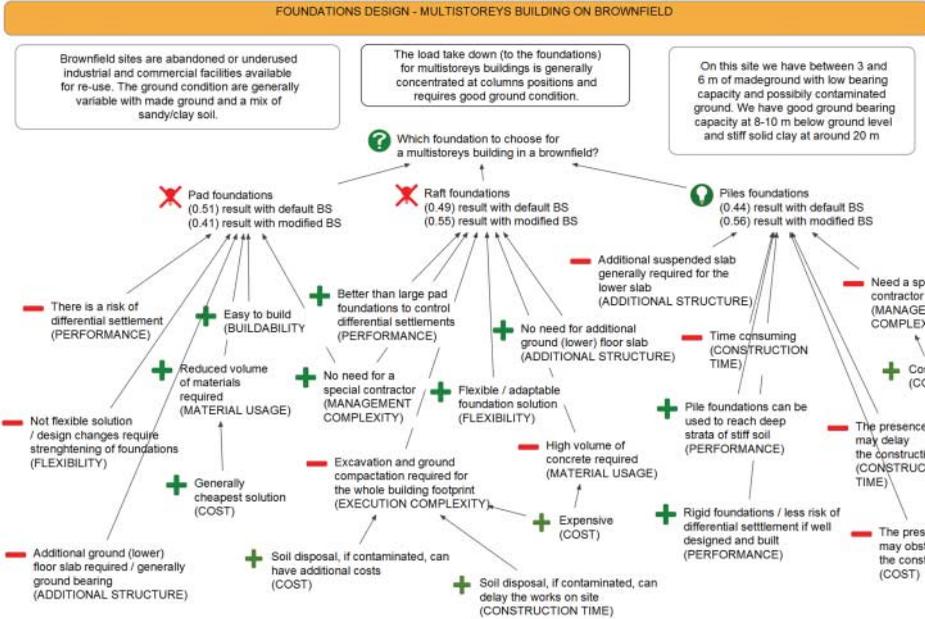


Figure 5. designVUE graph of the foundation project debate. Note that in designVUE answer nodes may have multiple (manually set) statuses (as in the original IBIS). In agreement with the automatic evaluation, the status for the Pad and Raft foundation answers has been manually changed to ‘rejected’ (red crossed out light bulb icon), while that for the Piles foundation answer to ‘accepted’ (green light bulb icon).

and managing explicitly quantitative valuations enhances transparency and accountability of the decision process.

7.2. Sludge reuse

This case concerns the selection of a technology for reuse of sewage sludge produced from the treatment of wastewater. Similarly to the case study considered in Section 7.1, this problem is well known and the relevant solution principles well established. Moreover, it is a real application example from previous experience of the collaborating expert. Two differences can be pointed out with respect to the case study in Section 7.1. First, the expert involved in the Sludge Reuse case had neither previous knowledge of the IBIS concept, nor of any tool implementing it. Second, the solution assessment is a two-step process, with different actors involved in each step, as described below.

Land application (A.1) has been the traditional sludge reuse option, due to its content of organic carbon and nutrients. Given that reuse in agriculture is subject to restrictions (since the sludge also contains pollutants), other disposal routes are considered as viable alternatives, such as reuse in the cement industry (A.2), energy recovery by combustion (A.3) or wet oxidation (A.4). The choice of the best alternative depends on technical (feasibility, applicability, reliability, etc.), economic, environmental and social factors (as pointed out by Achillas, Moussiopoulos, Karagiannidis, Baniyas, & Perkoulidis, 2013). In our case, nine factors were considered, five corresponding to pro-arguments (e.g. reliability) and four corresponding to con-arguments (e.g. vulnerability). While technical considerations, developed by experts, have been used to assign a score to each factor for each alternative, the importance of each factor cannot be established univocally, as it varies from site to site on the basis of other kinds of considerations. For instance,

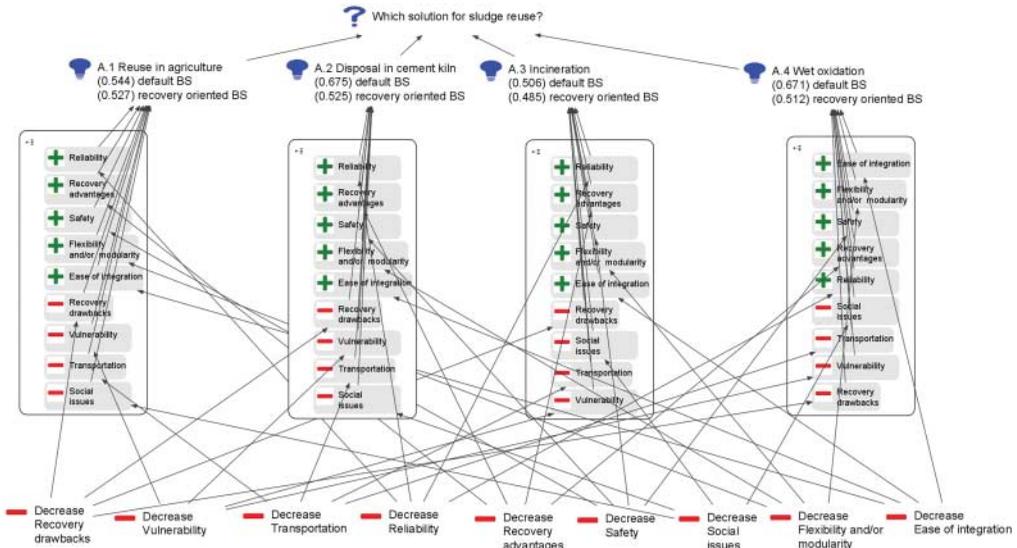


Figure 6. designVUE graph for sludge reuse technology selection. Note that the four answers are in the ‘open’ status (indicated as in IBIS by a blue light bulb icon) as the decision varies according to site-specific criteria.

the acceptability of a technology as perceived by the neighbouring population is a very important factor in an urban context (see the NIMBY (‘not in my back yard’) syndrome), while it is almost negligible for isolated locations. Hence, the final decision pertains to public officers or committees, who, taking into account context-specific aspects (e.g. social issues), may ascribe different importance to the various factors. To represent this two-phase decision process within designVUE, the expert suggested the use of a graph with a characteristic two-tier structure (Figure 6), where:

- the first tier takes into account the technical strengths and weaknesses of every single alternative. These are the pro- and con-arguments directly linked with the answers, whose base scores have been provided by the domain expert;
- each pro- or con-argument in the first tier has been assigned a weight ranging from 0 (for a factor which is irrelevant to a given alternative) to 0.1 (for a factor which is fully relevant to an alternative). For instance, the base score of the con-argument ‘Vulnerability’ linked to A.3 is 0 since combustion is deemed not to be vulnerable at all and is 0.1 for the corresponding con-argument linked to A.1 since land application is extremely vulnerable (e.g. to norm changes). Reuse in the cement industry A.2 has an intermediate degree of vulnerability (base score 0.05), while wet oxidation A.4 is not vulnerable at all (base score 0). The restriction of the range of the base scores corresponds to the choice, valid in this specific context, that each factor may individually affect the final score only to a limited extent: the higher the base score, the more a single factor can possibly play a saturating role. This individual saturating behaviour was deemed not appropriate in this case;
- the second tier pertains to the final decision-makers and consists of con-arguments against the pro- and con-arguments in the first tier. By assigning the base scores to the arguments of the second tier, the final decision-makers may modulate the actual influence of first-tier arguments according to context-specific considerations. The default base score of the second-tier arguments is 0, which corresponds to leaving the base scores assessed by

experts unaffected and to ascribe the same importance to all factors. The importance of each factor can be reduced by raising the base score of the corresponding con-argument in the second tier. The graph structure ensures that the same factor gets the same weight in the assessment of all alternatives.

Following this line, designVUE can be used to support a multistep methodology taking explicitly into account different classes of stakeholders. While the study of this methodology is left to future work, we carried out some preliminary experiments comparing the results obtained by varying the base scores of the second-tier arguments to show different attitudes towards the factors represented by the first-tier arguments. For instance, as shown in Figure 6, if all factors are deemed to have the same importance (i.e. the base score of all second-tier arguments is 0) then reuse in the cement industry is the preferred solution (with a final score of 0.675), followed by wet oxidation (0.671), land application (0.544) and combustion (0.506). Considering instead a scenario of strong preference for resource recovery-related factors, where the base score of the second-tier con-arguments attacking first-tier arguments not related to reuse is raised to 0.9, a different ranking is obtained where land application (final score 0.527) is the preferred solution, followed by reuse in the cement industry (0.525), wet oxidation (0.512) and combustion (0.485). This ranking is in accordance with the expert's expectation for this scenario: agriculture application is in effect the solution which allows complete material recovery (which is at high level in waste management hierarchy, accordingly with EU policies); on the contrary, combustion leads only to energy recovery, which is considered to be less valuable from an environmental perspective; reuse in the cement industry and wet oxidation processing can be considered as intermediate solutions, where energy recovery is predominant on material recovery.

As evidenced above, the development of this case study required several domain-specific modelling choices and, in fact, pointed out several open issues, first of all the need of methodological guidelines for the use of the formalism and the assessment of base scores. Nevertheless, the expert expressed a positive judgement about the results of the preliminary experiments carried out with the tool and a particular appreciation for the intuitive visual representation and the traceability of the reasons underlying the final decisions. He also remarked that in the environmental field a combination of qualitative and quantitative assessments is often used in decision processes and suggested that providing a formal counterpart to these hybrid evaluations is an important direction of future extension.

7.3. *Outpatient syringe*

This case study concerns the development of a syringe and is based on a design task reported in the literature (Ulrich & Eppinger, 2004) to illustrate concept selection by means of a well-known design method such as the decision matrix (Pugh, 1991). In particular, it compares our enhanced version of designVUE to an application of decision matrices for concept screening. Concept screening consists of making a first cut of the concepts proposed to solve a problem with a view to identifying those upon which to undertake refinement and scoring. The data used to populate this case were extracted from the decision matrix and other design information available in Ulrich & Eppinger (2004). The design problem entails choosing the best concept for an improved reusable syringe with precise dosage control for outpatient use (Ulrich & Eppinger, 2004). The problem is described in the matrix in Figure 7 where it can be seen that seven concepts (labelled A–G) were proposed, namely master cylinder, rubber brake, ratchet, plunge stop, swash ring, lever set and dial screw. Seven selection criteria, listed in the first column of the matrix, were identified to guide the decision. The upper part of the matrix was then filled in by carrying out a qualitative comparison of each concept against a reference solution (REF) for a given criterion. The outcome

SELECTION CRITERIA	CONCEPT VARIANTS							
	A	B	C	D	E	F	G	REF.
Ease of Handling	0	0	-	0	0	-	-	0
Ease of Use	0	-	-	0	0	+	0	0
Number Readability	0	0	+	0	+	0	+	0
Dose Metering	+	+	+	+	+	0	+	0
Load Handling	0	0	0	0	0	+	0	0
Manufacturing Ease	+	-	-	0	0	-	0	0
Portability	+	+	-	-	0	-	-	0
PLUSES	3	2	2	1	2	2	2	
SAMES	4	3	1	5	5	2	3	
MINUSES	0	2	4	1	0	3	2	
NET	3	0	-2	0	2	-1	0	
RANK	1	3	7	5	2	6	4	
CONTINUE?	Yes	Yes	No	No	Yes	No	Yes	

Figure 7. Matrix representing the outpatient syringe decision problem, from Ulrich & Eppinger (2004).

of the comparison is +, - or 0, meaning, respectively, that the concept is superior, inferior or equivalent to the reference as far as the criterion is concerned. These detailed evaluations are then summarised in the self-explaining lower part of the matrix.

The ranking in the penultimate row of the matrix in Figure 7 suggests, in particular, that the master cylinder concept (A) is preferable to all the others. It has to be observed, however, that the ranking of the rubber brake (B), plunge stop (D) and dial screw (G) has no explicit justification as they all have the same net score.

Figure 8 provides the representation of this problem in designVUE, with the nodes labelled with the strength *S* as computed through our enhanced version of designVUE, as well as the ranking given in Figure 7 for convenience. In absence of any indication, we have used the default base score of 0.5 for all the arguments. The results indicate that the master cylinder/A (with strength 0.93) is the preferred solution followed by the swash ring (0.87), the rubber brake, plunge stop and dial screw (0.5), the lever set (0.46) and the ratchet (0.45). The order of the ranking is largely in agreement with the matrix except for the rubber brake, plunge stop and dial screw. Indeed, more coherently with the available information, in the IBIS map these concepts get the

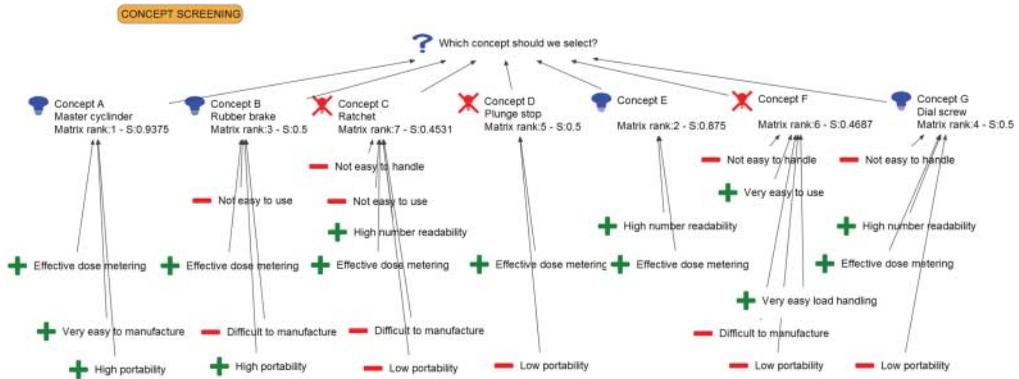


Figure 8. A designVUE representation of the problem in Figure 7. Here, answer nodes are equipped with the final score computed by the enhanced designVUE as well as the ranking in Figure 7.

same rank. Of course if the different ranking in the matrix is induced by some a priori preference or different weighting of the pros and cons, this can be encompassed in our approach using different base scores. It is noteworthy that in this case the results reflect the idea of counting pros and cons. More precisely it can be seen that for any argument a such that both attackers and supporters are present (namely $\mathcal{R}^-(a) \neq \{\}$ and $\mathcal{R}^+(a) \neq \{\}$), it holds that $\mathcal{SF}(a) = 0.5 + 0.5^{|\mathcal{R}^-(a)|} - 0.5^{|\mathcal{R}^+(a)|}$, where $|S|$ stands for the cardinality of S . Note that the higher the number of attackers the smaller the positive term is, while the higher the number of supporters the smaller the negative term is. For small numbers of attackers and supporters, the order of the final scores corresponds to the difference between the cardinalities of pros and cons, but it has to be noted that in our approach this difference has lesser effect with the increase of the number of attackers and supporters (e.g. having 12 supporters against 10 attackers give rise to a lesser final score than having 3 supporters against 1 attacker).

8. Related work

In engineering design, various methods are used to support the evaluation of design alternatives, for example, decision matrix (Pugh, 1991) and analytic hierarchy process (Saaty, 1980). Among these, the decision matrix, also known as the Pugh method, is the simplest and most commonly adopted. It consists of ranking alternatives by identifying a set of evaluation criteria, weighting their importance, scoring the alternatives against each criteria, multiplying the scores by the weight and computing the total score for each alternative (see our third case study in Section 7.3). Our work differs from the Pugh method in that it aims to extract a quantitative evaluation of alternatives from rich and explicitly captured argumentation rather than systematically assigned and justified scores. Hence, it seems to have the potential to lead to more logically reasoned decisions, as we discussed in Section 7.3.

The use of argumentation-based techniques has been advocated in several works in the engineering design literature.

The ABEN framework (Jin & Geslin, 2009) provides a detailed argumentation-based model of dialogues for a form of collaborative engineering design called *co-construction*. It encompasses protocols, strategies and tactics, but does not include any argument evaluation mechanism.

The DEEPFLOW project (Browne et al., 2011) aims at the extraction of formal arguments from design documents in natural language. This approach lies at a different modelling level than ours as it uses a logic-based argument representation rather than abstract argumentation frameworks. The paper exemplifies the use of probabilistic argumentation in this context without analysing the underlying mechanism in detail.

The approach of Liu, Raorane, Zheng, & Leu (2006) is more similar to ours. It models engineering design debates through *dialog graphs* featuring an IBIS-like structure with attack and support relations and argument weights in the $[-1, 1]$ interval. The dialog graph has a tree structure which is reduced to a one-layer tree (basically each argument is attached directly to the relevant answer) with modified weights using some heuristic fuzzy rules and a fuzzy set representation of the five possible qualitative interactions considered (strong/medium attack, indifference and strong/medium support). Thus, differently from our approach, a final score is produced only for answers, not for pro- and con-arguments. Formal properties of the proposed evaluation mechanism are not analysed by Liu, Raorane, Zheng, & Leu (2006); however, it can be observed that the behaviour of their approach heavily relies on the (somehow arbitrary) choice of the qualitative interaction and of the membership functions of the fuzzy sets representing them. In particular, it can be noted that, as evidenced by one of the examples presented by Liu, Raorane, Zheng, & Leu

(2006), some arguments with non-zero weight may turn out to have no impact on the final result since the inference mechanism produces the same result as if they were not present.

The HERMES system (Karacapilidis & Papadias, 2001), as well as its predecessor ZENO (Brewka & Gordon, 1994; Gordon & Karacapilidis, 1997), add numerical weights and constraints representing preferences to the basic elements of the IBIS model, giving rise to a hybrid quantitative/qualitative evaluation system. While considering hybrid evaluations is an interesting direction of future work, we remark that the use of numerical weights in HERMES is quite different from ours. Initial weights are first used to determine the so-called *activation* of arguments only in the case the proof standard called *scintilla of evidence* is adopted: an argument is active simply when the sum of the weights of its supporters is greater than the sum of the weights of its attackers. The subsequent phases of the argumentation process do not use weights that come back into play in the final stage, where, for each alternative, a minimum and a maximum weight compatible with the constraints is computed and the final weight of the alternative is their average.

The CoPe_it! system (Karacapilidis et al., 2009) uses a similar argument evaluation method as HERMES within an enriched, web-based environment for visualisation of debates, providing users with means to organise and structure data as well as import legacy resources.

Also the EDEN system (Marashi & Davis, 2006) provides the visualisation and the automated evaluation of engineering design debates using an IBIS-like model. The numerical evaluation mechanism is based on a variation of Dempster–Shafer theory of evidence and produces a pair of values, called belief and plausibility, for each argument. While a detailed analysis of this approach is outside the scope of this paper, we remark the basic difference that evidence theory deals with uncertainty quantifications while our approach concerns a notion of gradual acceptability, which is, conceptually, an orthogonal dimension with respect to uncertainty.

None of the above-mentioned papers includes a detailed analysis of the basic properties of the proposed numerical formalism nor of the relationships with non-numerical formalisms of the kind provided in Sections 4 and 5.

Our system extends an existing IBIS-based tool, designVUE, already used in the engineering domain and in particular familiar to some of the experts responsible for our case studies. Other IBIS-based systems exist in the literature. For example, Cohere and Compendium (Buckingham Shum, 2008; Buckingham Shum et al., 2006), adopt an IBIS methodology to support design rationale in collaborative settings. However, these systems do not incorporate means to automatically evaluate debates. Other examples are the Carneades (Gordon & Walton, 2006) and the PARMENIDES (Atkinson, Bench-Capon, & McBurney, 2006) systems. These adopt a more articulate model of debate as they use argument schemes and critical questions as basic building blocks of the argumentation process. However, they do not incorporate a numerical evaluation of positions in debates. The extension of these other systems to take advantage of our scoring methodology is a possible direction of future work.

Turning to argumentation literature, the idea of providing a quantitative evaluation of a given position on the basis of arguments in favour and against has been considered in several works.

In Besnard & Hunter (2001), in the context of a logic-based approach to argumentation, an argument structure for a logical formula α is (omitting some details) a collection of reasons supporting $(\neg)\alpha$. Each reason is represented as an argument tree, whose root is an argument for $(\neg)\alpha$ and where the children of an argument node are attackers of the node itself. Each argument tree is quantitatively evaluated using a *categoriser*. The results of the evaluation of argument trees for $(\neg)\alpha$ are aggregated separately using an *accumulator* function and then combined. Though this work shows several similarities with our approach at a generic level, we point out some important differences. In Besnard & Hunter (2001), the evaluation concerns logical formulas rather than arguments, arguments can only attack (not support) each other, while the notion of

support for a formula coincides with the (defeasible) derivation of the formula. Then, differently from our approach, the recursive procedure corresponding to the categoriser concerns attacks only and the notion of support plays a role only in the accumulator. Also, in Besnard & Hunter (2001), there is no notion of base score.

The gradual valuation of BAFS (Cayrol & Lagasque-Schiex, 2005b) (Section 1.2) is closer to our proposal. In fact, the generic valuation function ν of BAFS (Definition 1.5) has a similar structure to our \mathcal{SF} , with h_{sup} , h_{def} corresponding to our \mathcal{F}_{supp} , \mathcal{F}_{att} , respectively, and satisfying analogous properties. A basic difference concerns the base score, absent in Cayrol & Lagasque-Schiex (2005b) and crucial in our application domain.

The ESAAF approach of Evripidou & Toni (2012) (Section 1.2) has more similarities, as it encompasses an initial score for arguments (obtained from votes) and a recursive evaluation mechanism similar to ours. In fact, the treatment we propose for attackers coincides with the one proposed in Evripidou & Toni (2012), while our proposal differs in the treatment of supporters: in Evripidou & Toni (2012), supporters are treated as a sort of ‘negative attacks’, while in our approach supporters contribute to increase the base score specularly to the way attackers contribute to decrease it. As a consequence, in ESAAF the operator used for the combination of the initial score with the aggregation of supporters’ valuations includes the min operator to prevent that the combination exceeds the limit value of 1. This means that the contribution of supporters is subject to a saturation which may be undesirable in some cases. This difference is made more obvious by the reformulation in Evripidou & Toni (2014) of the ESAAF approach in the same style as our approach in this paper, but using variants of g , \mathcal{F}_{att} , \mathcal{F}_{supp} that we use, so as to obtain an equivalent re-interpretation of the original ESAAF method that can be more easily compared with the QuAD framework.

The approach of Gabbay (2012) also features significant similarities with our proposal. In fact the notion of *real equational network* introduced in Gabbay (2012) uses an evaluation function $f(a)$ from the set of arguments to $[0, 1]$ which is defined recursively, for an argument a , as $f(a) = h_a(f(a_1), \dots, f(a_k))$ where a_1, \dots, a_k are the attackers of a . Gabbay (2012) explores several alternatives for the function f with unrestricted graph topology (in the presence of cycles, the solution is a fixed point of f) but no notion of the base argument score is considered. Note that, assuming a fixed base score of 1 for any argument, our \mathcal{F}_{att} coincides with the function called *Eqinverse* in Gabbay (2012). Gabbay (2012) considers also the presence of a support relation, but treated as a potential ‘vehicle’ for attacks, in the sense that if an argument a supports another argument b , an attacker of a is also considered as an (indirect) attacker of b and contributes to decreasing its score. On the other hand, a supporting argument cannot increase the score of the supported argument. This view is coherent with the absence of a base score and is clearly alternative to ours.

Other approaches to quantitative valuation have been proposed in the context of Dung’s abstract argumentation where only the attack relation is encompassed. For example, Matt & Toni (2008) propose a game-theoretic approach to evaluate argument strength in abstract argumentation frameworks. In a nutshell, the strength of an argument x is the value of a game of argumentation strategy played by the proponent of x . The approach does not encompass support relations nor base scores: extending this game-theoretic perspective with these notions appears to be a significant direction of future investigation. Also, in weighted argumentation frameworks (Dunne, Hunter, McBurney, Parsons, & Wooldridge, 2011), real-valued weights are assigned to attacks (rather than to arguments). These weights are not meant to be a basis for scoring arguments, rather they represent the ‘amount of inconsistency’ carried by an attack. This use of weights is clearly different from ours and, in a sense, complementary. Investigating a combination of these two kinds of valuations (possibly considering also weights for support links) is a further interesting direction of future work.

9. Conclusions

We presented a novel argumentation-based formal framework for quantitative assessment of design alternatives, its implementation in the designVUE software tool and its preliminary experimentation in three case studies. In addition to those mentioned in Section 8, several directions of future work can be considered. On the theoretical side, a more extensive analysis of the properties of the proposed score function is under way, along with the study of alternative score functions exhibiting a different behaviour (e.g. concerning the effect of attackers and supporters and their balance) while satisfying the same basic requirements. On the implementation side, we plan to integrate the QuAD framework in a web-based debate system similar to www.quaestio-it.com so to gain experience on its acceptability by users in other domains. On the experimentation side, the development of further engineering design case studies (more complex and/or in other domains) is under way and we intend to continue with the on-field comparison with more traditional approaches to the evaluation of design alternatives that we have started with the third case study in this paper.

Acknowledgements

The authors are grateful to the anonymous reviewers for their helpful comments. The authors thank V. Evripidou and E. Marfisi for their support and cooperation. Aurisicchio and Toni thank the support of a Faculty of Engineering EPSRC Internal Project on ‘Engineering design knowledge capture and feedback’.

Conflict of interest disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Here, separability amounts to absence of interaction between attackers and supporters.
2. The expression of f_{supp} corresponds to the T-conorm operator also referred to as *probabilistic sum* in the literature (Klement, Mesiar, & Pap, 2000).
3. www3.imperial.ac.uk/designengineering/tools/designvue.
4. The code is available from the designVUE web site.

ORCID

Pietro Baroni  <http://orcid.org/0000-0001-5439-9561>

Francesca Toni  <http://orcid.org/0000-0001-8194-1459>

Marco Aurisicchio  <http://orcid.org/0000-0002-1119-4336>

Giorgio Bertanza  <http://orcid.org/0000-0002-2965-023X>

References

- Achillas, C., Moussiopoulos, N., Karagiannidis, A., Baniyas, G., & Perkoulidis, G. (2013). The use of multi-criteria decision analysis to tackle waste management problems: A literature review. *Waste Management & Research*, 31, 115–129.
- Amgoud, L., Cayrol, C., Lagasque-Schiex, M. C., & Livet, P. (2008). On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23, 1062–1093.
- Atkinson, K., Bench-Capon, T. J. M., & McBurney, P. (2006). PARMENIDES: Facilitating deliberation in democracies. *Artificial Intelligence and Law*, 14, 261–275.

- Auriscchio, M., & Bracewell, R. (2013). Capturing an integrated design information space with a diagram-based approach. *Journal of Engineering Design*, 24, 397–428.
- Baroni, P., Caminada, M., & Giacomin, M. (2011). An introduction to argumentation semantics. *The Knowledge Engineering Review*, 26, 365–410.
- Baroni, P., Romano, M., Toni, F., Auriscchio, M., & Bertanza, G. (2013, September 16–18). An argumentation-based approach for automatic evaluation of design debates. In J. Leite, T. C. Son, P. Torroni, L. van der Torre, & S. Woltran (Eds.), *Computational Logic in Multi-Agent Systems, 14th International Workshop, CLIMA XIV, Corunna, Spain, Proceedings*, Lecture Notes in Artificial Intelligence (Vol. 8143, pp. 340–356). Berlin: Springer-Verlag.
- Besnard, P., & Hunter, A. (2001). A logic-based theory of deductive arguments. *Artificial Intelligence*, 128, 203–235.
- Brewka, G., & Gordon, T. (1994, July 31). How to buy a Porsche: An approach to defeasible decision making. In *Working notes of the AAAI-94 Workshop on Computational Dialectics* (pp. 28–38). Seattle, WA: AAAI Press.
- Browne, F., Jin, Y., Higgins, C., Bell, D., Rooney, N., Wang, H., . . . Taylor, P. (2011). *The application of a natural language argumentation based approach within project life cycle management*. Proceedings of the 22nd Irish Conference on Artificial Intelligence and Cognitive Science, University of Ulster, Magee Campus, Derry, Northern Ireland.
- Buckingham Shum, S. J. (2008, May 28–30). Cohere: Towards Web 2.0 argumentation. In P. Besnard, S. Doutre, & A. Hunter (Eds.), *Proceedings of the 2nd International Conference on Computational Models of Argument (COMMA 2008)*, Toulouse, France, Vol. 172 of Frontiers in Artificial Intelligence and Applications (pp. 97–108). Amsterdam: IOS Press.
- Buckingham Shum, S. J., & Hammond, N. (1994). Argumentation-based design rationale: What use at what cost? *International Journal of Human-Computer Studies*, 40, 603–652.
- Buckingham Shum, S. J., Selvin, A. M., Sierhuis, M., Conklin, J., Haley, C. B., & Nuseibeh, B. (2006). Hypermedia support for argumentation-based rationale: 15 Years on from gIBIS and QOC. In A. H. Dutoit, R. McCall, I. Mistrik, & B. Paech (Eds.), *Rationale management in software engineering* (pp. 111–132). Berlin: Springer.
- Cayrol, C., & Lagasque-Schiex, M. C. (2005a, July 6–8). On the acceptability of arguments in bipolar argumentation frameworks. In L. Godo (Ed.), *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, Barcelona, Spain, Lecture Notes in Computer Science (Vol. 3571, pp. 378–389). Springer.
- Cayrol, C., & Lagasque-Schiex, M. C. (2005b, July 6–8). Gradual valuation for bipolar argumentation frameworks. In L. Godo (Ed.), *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2005)*, Barcelona, Spain, Lecture Notes in Computer Science (Vol. 3571, pp. 366–377). Berlin: Springer.
- Cayrol, C., & Lagasque-Schiex, M.-C. (2013). Bipolarity in argumentation graphs: Towards a better understanding. *International Journal of Approximate Reasoning*, 54, 876–899.
- 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.
- Dunne, P. E., Hunter, A., McBurney, P., Parsons, S., & Wooldridge, M. (2011). Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175, 457–486.
- Evrpidou, V., & Toni, F. (2012, September 10–12). Argumentation and voting for an intelligent user empowering business directory on the web. In M. Krötzsch & U. Straccia (Eds.), *Proceedings of the 6th International Conference on Web Reasoning and Rule Systems (RR'12)*, Vienna, Austria. Lecture Notes in Computer Science (Vol. 7497, pp. 209–212). Berlin: Springer.
- Evrpidou, V., & Toni, F. (2014). Quaestio-it.com: A social intelligent debating platform. *Journal of Decision Systems*, 23, 333–349.
- Fischer, G., Lemke, A. C., McCall, R., & Morch, A. I. (1991). Making argumentation serve design. *Human-Computer Interaction*, 6, 393–419.
- Gabbay, D. M. (2012). Equational approach to argumentation networks. *Argument & Computation*, 3, 87–142.

- Gordon, T. F., & Karacapilidis, N. I. (1997, June 30–July 3). The Zeno Argumentation Framework. In *Proceedings of the 6th International Conference on Artificial Intelligence and Law (ICAIL'97)* (pp. 10–18). Melbourne, Victoria, Australia.
- Gordon, T. F., & Walton, D. (2006, September 11–12). The Carneades argumentation framework – using presumptions and exceptions to model critical questions. In P. E. Dunne & T. J. M. Bench-Capon (Eds.), *Proceedings of the 1st International Conference on Computational Models of Argument (COMMA 2006)*, Liverpool, UK. *Frontiers in Artificial Intelligence and Applications* (Vol. 144, pp. 195–207). Amsterdam: IOS Press.
- Jin, Y., & Geslin, M. (2009). Argumentation-based negotiation for collaborative engineering design. *International Journal of Collaborative Engineering*, 1, 125–151.
- Karacapilidis, N., & Papadias, D. (2001). Computer supported argumentation and collaborative decision making: The HERMES system. *Information Systems*, 26, 259–277.
- Karacapilidis, N., Tzagarakis, M., Karousos, N., Gkotsis, G., Kallistros, V., Christodoulou, S., & Mettouris, C. (2009). Tackling cognitively-complex collaboration with CoPe_it! *International Journal of Web-based Learning and Teaching Technologies*, 4, 22–38.
- Klement, E. P., Mesiar, R., & Pap, E. (2000). *Triangular norms*. Dordrecht: Kluwer.
- Kunz, W., & Rittel, H. (1970). *Issues as elements of information systems* (Working Paper 131). Berkeley, CA: Institute of Urban and Regional Development, University of California.
- Leite, J., & Martins, J. (2011, July 16–22). Social abstract argumentation. In T. Walsh (Ed.), *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*, Barcelona, Catalonia, Spain (pp. 2287–2292). Menlo Park, CA: AAAI Press/International Joint Conferences on Artificial Intelligence.
- Liu, X., Raorane, S., Zheng, M., & Leu, M. (2006, May 14–17). An Internet based intelligent argumentation system for collaborative engineering design. In W. W. Smari & W. K. McQuay (Eds.), *Proceedings of CTS 2006, International Symposium on Collaborative Technologies and Systems*, Las Vegas, NV, USA (pp. 318–325) Los Alamitos, CA: IEEE Computer Society.
- Marashi, E., & Davis, J. P. (2006). An argumentation-based method for managing complex issues in design of infrastructural systems. *Reliability Engineering & System Safety*, 91, 1535–1545.
- Matt, P. A., & Toni, F. (2008, September 28–October 1). A game-theoretic measure of argument strength for abstract argumentation. In S. Hölldobler, C. Lutz, & H. Wansing (Eds.), *Proceedings of the 11th European Conference on Logics in Artificial Intelligence (JELIA 2008)*, Dresden, Germany. *Lecture Notes in Computer Science* (Vol. 5293, pp. 285–297). Berlin: Springer.
- Pahl, G., & Beitz, W. (1984). *Engineering design: A systematic approach* (Technical report). London: Design Council.
- Pugh, S. (1991). *Total design: Integrated methods for successful product engineering*. Boston, MA: Addison-Wesley.
- Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. New York: McGraw-Hill.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
- Simon, H. A., & Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist*, 26, 145–159.
- Ulrich, K. T., & Eppinger, S. D. (2004). *Product design and development* (3rd ed.). New York: Irwin McGraw-Hill.