Context-based and Explainable Decision Making with Argumentation

Zhiwei Zeng†
LILY Research Centre, IGS
Nanyang Technological University
Alibaba-NTU Singapore JRI
i160001@e.ntu.edu.sg

Xiuyi Fan‡
LILY Research Centre, NTU
Department of Computer Science
Swansea University
xiuyi.fan@swansea.ac.uk

Chunyan Miao‡
LILY Research Centre, SCSE
Nanyang Technological University

Cyril Leung
LILY Research Centre
Nanyang Technological University

Chin Jing Jih§
LKC School of Medicine, NTU
Centre for Geriatric Medicine, TTSH

Ong Yew Soon§
SCSE
Nanyang Technological University

ABSTRACT
Argumentation-based approaches to decision making have gained considerable research interest, due to their ability to select and justify decisions. In order to make better decisions, context is a key piece of information that needs to be considered. However, most existing argumentation-based models and frameworks have not modelled or reasoned with context explicitly. In this paper, we present a new argumentation-based approach for making context-based and explainable decisions. We propose a graphical representation for modelling decision problems involving varying contexts, Decision Graphs with Context (DGC), and a reasoning mechanism for making context-based decisions which relies on the Assumption-based Argumentation formalism. Based on these constructs, we introduce two types of explanations, argument explanation and context explanation, identifying the reasons for the decisions made from an argument-view and a context-view respectively.

KEYWORDS
Decision making; context-awareness; argumentation; explanation

ACM Reference Format:

1 INTRODUCTION
Amongst various approaches to decision making, argumentation-based approaches have gained increasing amount of research interest recently [1, 14, 21, 32]. Argumentation can play two different roles in decision making, namely help to select, or to explain and justify decisions. Argumentation-based approaches to decision making are expected to be more akin with the way humans deliberate, evaluate alternatives and make decisions [2]. This endows argumentation-based approaches with unique benefits, including transparent decision making process and the ability to offer understandable reasons underlying the decisions made.

Context, the particular situation, environment or domain in which a decision is to be made, is a key piece of information that needs to be taken into account in order to make an optimal decision. Contexts add additional dynamics and complexity to decision making in a sense that a decision may be “good” in a particular context but less “good” in other contexts. Studies in [4] have shown that humans evaluate arguments differently depending on implicit domain knowledge. Hence, incorporating context in problem modelling and reasoning can help to establish better understanding of the decision parameters and make holistic evaluation of the decision alternatives. However, existing argumentation-based approaches to decision making have not modelled or reasoned with context explicitly during the decision making process, making it hard to study and understand the effects of context on decisions.

In this paper, we present an argumentation-based approach for making context-based and explainable decisions. To introduce context into decision making, the first step is to model context in the formal representation of a decision problem. We propose Decision Graphs with Context (DGC) for this purpose. DGCs can capture the varying relationships between decisions and goals in different contexts, offering greater expressiveness and flexibility in modelling decision problems. To select “good” decisions, we map DGCs to Assumption-based Argumentation (ABA) frameworks and transform the process of making context-based decisions in DGCs to determining argument admissibility in ABA frameworks.

To make the decision making process more transparent to humans, we propose two types of explanations for the decisions made. It is useful to study the reasons for not selecting a decision for the purpose of improving the decision alternatives or adapting decisions in different contexts. We introduce argument explanations to explain the sources of failure of a decision alternative by identifying the attackers from which it cannot be defended. Sometimes an alternative is not “good” in a given context but “good” in another. Benefiting from the new graphical representation and the reasoning
mechanism proposed, we also introduce context explanations which give context-specific reasons to explain whether the failure of a decision can be attributed to contexts.

The remaining part of this paper is organised as follows. We introduce relevant background in Section 2 and present Decision Graphs with Context (DGC) in Section 3. We then illustrate how to compute "good" decisions with Assumption-Based Argumentation (ABA) in Section 4 and how to derive the two forms of explanations in Section 5. Finally, we discuss related works in Section 6 and conclude in Section 7.

2 BACKGROUND

Abstract Argumentation (AA) frameworks [8] are pairs $AF = \langle B, K \rangle$, consisting of a set of arguments, $B$, and a binary attack relation, $K$. Given an AA framework $AF = \langle B, K \rangle$, a set of arguments $B \subseteq B$ is admissible in $AF$ iff $\forall a, b \in B$, there exists no $(a, b) \in K$ (if is conflict free) and $\forall a \in B$, if $(c, a) \in K$, then there exists some $b \in B$ such that $(b, c) \in K$.

We say that an argument $a$ is in $AF$ iff $a \in B$, and an attack $(a, b)$ is in $AF$ or $a$ attacks $b$ in $AF$ iff $(a, b) \in K$.

Assumption-based Argumentation (ABA) frameworks [29] are tuples $(L, R, A, C)$ where

- $(L, R)$ is a deductive system, with a language $L$ and a rule set $R$ of the form $R_0 \leftarrow R_1, \ldots, R_m (m \geq 0, R_i \in L)$;
- $A \subseteq L$ is a non-empty set, referred to as assumptions;
- $C$ is a total mapping from $A$ into $2^L$, where each $c \in C(a)$ is a contrary of $a$.

Given a rule $\rho$ of the form $R_0 \leftarrow R_1, \ldots, R_m, R_0$ is referred to as the head and $R_1, \ldots, R_m$ as the body of $\rho$. All ABA frameworks are flat, i.e. assumptions do not occur in the head of rules.

In ABA frameworks, arguments are deductions of claims with sets of rules and supported by sets of assumptions. Attacks against arguments are directed at the assumptions in the support of arguments. Informally, adapted from [9, 29]:

- an argument for $\beta \in L$ supported by $\Delta \subseteq A$ with $R \subseteq R$ (denoted $\Delta \vdash R \beta$) is a finite tree with nodes labelled by sentences in $L$ or by $\top$;
- the root labelled by $\beta$, leaves either $\top$ or assumptions in $A$, and non-leaves $\beta'$ with the elements of the body of some rule in $R$ with head $\beta''$ as children, and $R$ contains no other rules except the ones in the tree;
- an argument $\Delta_1 \vdash R_1, \beta_1$ attacks an argument $\Delta_2 \vdash R_2, \beta_2$ iff $\beta_1$ is a contrary of one of the assumptions in $\Delta_2$.

When there is no ambiguity, $\Delta \vdash \beta$ is used as the shorthand form for $\Delta \vdash R \beta$.

Admissibility and other semantics introduced in AA can also be applied to ABA [12]. Formally, given an ABA framework $ABF = \langle L, R, A, C \rangle$, a set of assumptions is admissible in ABA iff it does not attack itself and it attacks all $\Delta \subseteq A$ that attack it. We say that an argument $\Delta \vdash \beta$ is admissible in ABA iff there is an admissible set $\Delta' \subseteq A$ for which $\Delta \subseteq \Delta'$. We also say that an argument $\Delta \vdash R \beta$ is in ABA iff $R \subseteq R$ and $\Delta \subseteq A$.

Since ABA is an instance of AA, given an ABA framework, a corresponding AA framework can be constructed by following the procedures described in [10].

Explanations for non-admissible arguments in AA are defined using the pruning operator, \$. Given an AA framework $AF = \langle B, K \rangle$ and a set of arguments $B \subseteq B$, the repaired framework is $AF \setminus B = \langle B', K' \rangle$, where $B' = B \setminus B$ and $K' = \{((a, b)|(a, b) \in K \and a \in B', b \in B')\}$.

Given an AA framework $AF = \langle B, K \rangle$, let $a \in B$ be some non-admissible argument in $AF$. Then, $B \subseteq B$ is an explanation of $a$ iff: (1) $a$ is admissible in $AF \setminus B$, and (2) there exists no $B' \subseteq B$ such that $a$ is admissible in $AF \setminus B'$. If no such $B$ exists in $AF$, then $\{a\}$ is the explanation of $a$ [15].

A decision problem can be represented by a decision framework which describes the relationships between decisions and attributes and between goals and attributes with two tables, as follows:

A Decision Framework (DF) [11] is a tuple $(D, A, G, T_{DA}, T_{GA})$, consisting of:

- a finite set of decisions $D = \{d_1, \ldots, d_n\}, (n > 0)$,
- a finite set of attributes $A = \{a_1, \ldots, a_m\}, (m > 0)$,
- a finite set of goals $S = \{g_1, \ldots, g_l\}, (l > 0)$, and
- two tables $T_{DA}$, of size $n \times m$, and $T_{GA}$, of size $l \times m$, such that
  - for every $T_{DA}[i,j]$, $1 \leq i \leq n, 1 \leq j \leq m$, $T_{DA}[i,j]$ is either 1, representing $d_i$ has $a_j$, or 0, otherwise,
  - for every $T_{GA}[k,j]$, $1 \leq k \leq l, 1 \leq j \leq m$, $T_{GA}[k,j]$ is either 1, representing $g_k$ is satisfied by $a_j$, or 0, otherwise.

Given a decision framework $DF = \langle D, A, G, T_{DA}, T_{GA} \rangle$, a decision $d_i \in D$ meets a goal $g_k \in G$, with respect to $DF$, if and only if there exists an attribute $a_j \in A$, such that $T_{DA}[i,j] = 1$ and $T_{GA}[k,j] = 1$.

We use $\Gamma(d) = S$, where $d \in D, S \subseteq G$ to denote the set of goals met by $d$.  

3 MODELLING CONTEXT IN DECISION MAKING

We introduce Decision Graphs with Context (DGC) as a new representation for modelling decision problems.

A DGC contains two parts of information: (1) a directed acyclic graph with nodes and edges that represents the relationships between decisions and goals; (2) contexts in which the decision is to be made. In a DGC, there are three types of nodes, namely decisions, goals and intermediates, corresponding to the candidate decisions, goals and decision attributes. Edges represent relations amongst the nodes, e.g. an edge from a decision to an intermediate attribute represents that the decision possesses the attribute; an edge from an intermediate attribute to a goal represents that the attribute satisfies the goal; an edge from one intermediate attribute to another intermediate attribute represents the former leads to the latter. In a DGC, the relationship between two nodes can be either definite or defeasible. A definite relationship holds in all contexts while a defeasible relationship generally holds but becomes inapplicable in certain contexts. Formally:

Definition 3.1. A Decision Graph with Context (DGC) is a tuple $(N, E, C)$, in which $(N, E)$ is a directed acyclic graph, such that:

- $N = N_d \cup N_{int} \cup N_g$ is a set of nodes, such that $N_d, N_{int}$ and $N_g$ are pairwise disjoint, in which
  - $N_d \neq \emptyset$ is a set of decision nodes (decisions);
  - $N_{int}$ is a set of intermediate nodes (intermediates);
  - $N_g \neq \emptyset$ is a set of goal nodes (goals).
Session 27: Argumentation  

![Diagram](image)

**Figure 1:** A DGC Example. Solid arrows (→) represent strict edges $E_s$ whereas dashed arrows (→ ) represent defeasible edges $E_d$. Intuitive readings of the nodes in this graph are: clo stands for administer_clopidogrel, asp stands for administer_aspirin, dn means do nothing, rpa means reduced_platelet_adhesion, na means no_allergy, and se represents smaller_expense.

- $E = E_s \cup E_d$ is a set of edges, such that $E_s \cap E_d = \emptyset$, and
- $E_s$ is a set of strict edges;
- $E_d$ is a set of defeasible edges;
- $[n_i | n_j]$ denotes an edge from $n_i$ to $n_j$, for $n_i, n_j \in \mathbb{N}$;
- $[n_i | n_j]$ is in $E$ iff either $n_i \in N_d$ and $n_j \in N_{int} \cup N_g$, or $n_i \in N_{int}$ and $n_j \in N_{int} \cup N$;
- each edge $e \in E$ is associated with a tag $t$ such that $t(e) = i$.

- $C = C_p \cup C_r$ is a set of defeasible context such that $C_p \cap C_r = \emptyset$, in which
  - each sentence $c \in C$ is an implication of the form $t_n \wedge \ldots \wedge t_1 \rightarrow t_0 (n \geq 0)$ over a language $\mathcal{L}_c$ such that for each defeasible edge $e = [n_i | n_j'] \in E_d$ with $t(e) = i$ there is a defeasible context sentence $\neg dEdge(n, n', i) \in \mathcal{L}_c$;
  - $C_p$ is a set of context primitives;
  - $C_r$ is a set of rules;
  - for each $c \in C$ in the form $t_n \wedge \ldots \wedge t_1 \rightarrow t_0$, $n = 0$ if $c \in C_p$, and $n > 0$ if $c \in C_r$.

Given a DGC, strict edges in $E_s$ represent definite relationships between nodes while defeasible edges in $E_d$ capture defeasible relationships. $C_p$ consists of context primitives whereas $C_r$ contains rules that specify how the context primitives, either by themselves or in conjunction, would influence the defeasibility of an edge. The notion $\neg dEdge(n, n', i)$ represents that the defeasible edge $[n_i | n_j']$ is inapplicable (refer to Definition 3.3), and hence cannot be traversed. Tags associated with the edges allow more complex relationships to be captured and are used to determine whether a node is reachable (refer to Definition 3.5) from a set of nodes in a DGC.

**Example 3.2.** As a running example to illustrate our approach, we discuss the decision making problem of choosing the appropriate treatment for a patient threatened by blood clotting. This example is adapted from an example by Modgil on the treatment of heart disease [23].

There are three decision alternatives available for treating blood clotting: administer_clopidogrel (clo), administer_aspirin (asp) and simply do nothing (dn). The use of clopidogrel or aspirin can lead to reduced_platelet_adhesion (rpa), which reduces the risk of heart disease. The patient should have no_allergy (na) to the medicine administered in order to use it safely. In addition, if the patient has a history of gastritis, clopidogrel and aspirin should be used in conjunction with a proton pump inhibitor $^2$ to prevent gastrointestinal bleeding (which is considered unsafe) due to reduced platelet adhesion. In terms of cost considerations, aspirin is more affordable and incurs smaller_expense (se) than clopidogrel. Doing nothing incurs no monetary costs.

Specifically for Example 3.2, we consider a patient who has a history of gastritis and an allergy to clopidogrel. Meanwhile, inhibitors are in stock.

We construct the DGC for deciding the appropriate treatment as follows:
- the decisions are: $N_d = \{ \text{clo, asp, dn} \}$;
- the goals are: $N_g = \{ \text{safe, cheap} \}$;
- the intermediate attributes are: $N_{int} = \{ \text{rpa, na, se} \}$;
- the strict edges are: $E_s = \{ [\text{clo|rpa}], [\text{asp|rpa}], [\text{na|safe}], [\text{asp|se}], [\text{dn|se}], [\text{se|cheap}] \}$;
- the defeasible edges are:
  - $E_d = \{ [\text{clo|na}], [\text{asp|na}], [\text{rpa|safe}] \}$;
- the context primitives are: $C_p = \{ \text{gastritis, } \text{clo_allergy, } \text{inhibitor} \}$;
- the context rules $C_r$ are listed in Fig. 1.

The defeasible context $C = C_p \cup C_r$ specifies the contexts in which the defeasible edges become inapplicable and hence untraversable. In this example, the context primitives $C_p$ include medical history of the patient and the availability of the inhibitor. The context rules $C_r$ can be used to determine the defeasibility of $E_d$ given $C_p$. For instance, an intuitive reading of the context rule

$$\text{gastritis} \land \neg \text{inhibitor} \rightarrow \neg \text{dEdge(rpa, safe, 1)}$$

associated with edge $[\text{rpa|safe}]$ is: if the patient has a history of gastritis and no inhibitor can be used together with the medicine, reducing platelet adhesion does not achieve the goal safe.

Formally, with the defeasible context specified by $C$, we define inapplicable edges as follows.

**Definition 3.3.** Given a DGC $CG = \langle N, E, C \rangle$, with $E = E_s \cup E_d$ and $C = C_p \cup C_r$, the inapplicable edges of $CG$ is a subset of $E_d$, such that:

$$E_{IA} = \{ e \in E_d | e = [n_i | n_j], t(e) = k, C \vdash_{pp} \neg \text{dEdge}(n_i, n_j, k) \}$$

where $\vdash_{pp}$ stands for repeated applications of the modus ponens inference rule to the set of defeasible context $C$ until the elements of $E_{IA}$ do not change any more. $^3$

We use $\Psi(CG) = E_{IA}$ to denote the set of inapplicable edges in the context $C$.

**Example 3.4.** (Example 3.2 continued.) According to Definition 3.3, only defeasible edges can become inapplicable. With currently defined $C$, among the three defeasible edges $E_d = \{ [\text{clo|na}], [\text{asp|na}],$

$^2$There are still controversies regarding the concomitant use of clopidogrel and proton pump inhibitors.

$^3$The modus ponens inference rule amounts to deriving $c$ from either $\rightarrow c$ or $a \rightarrow c$ and $a$, for any set (conjunction) of sentences $a$ and sentences $c$.}
[rpa|safe] in the DGC in Fig. 1, only edge [clo|na] is inapplicable, i.e. $E_{ia} = \{[clo|na]\}$.

In order to provide means for determining whether a decision meets a goal, we first introduce the notion of reachability from a set of nodes to a node in DGC as follows.

**Definition 3.5.** Given a DGC $CG = (N, E, C)$, let $n \in N, N \subseteq N$. We say that $n$ is reachable from $N$ if and only if one of the following two conditions hold:

1. There exists a tag $k$ such that $N = \{n\}$ and $[n] \in E \setminus \Psi(CG)$ and $t(e) = k$; or
2. There exists some $N' \subseteq N$ such that $n$ is reachable from $N'$ and for each $n' \in N'$, $n'$ is reachable from $N$.

The reachability to a node is defined recursively. Item 1 specifies the base condition that a node $n$ is reachable from a set of nodes $N$. Item 2 specifies the "transitive" characteristic of reachability that if a node $n$ is reachable from some intermediate set $N'$ such that each node $n'$ in $N'$ is reachable from $N$, then $n$ is reachable from $N$.

Tags associated with the edges allow more complex relationships to be captured and are used when determining whether a node is reachable. A node $n$ is reachable from a set of nodes $N$ if all nodes in $N$ lead to $n$ via applicable edges labelled by the same tag. When the tag is 1, it is often omitted.

**Example 3.6.** In Fig. 1, there are two paths to reach node safe from node clo, one via node rpa and the other via node na. The two edges leading to node safe, [rpa|safe] and [na|safe], have the same tag 1 (both are 1, hence both are omitted). Since the two edges have the same tag, they have an "AND" relationship. The node safe is only reachable via the two paths simultaneously. Hence, node safe is reachable from [clo] only when both rpa and na are reachable from [clo] and safe is reachable from [rpa|na]. However, if the two edges, [rpa|safe] and [na|safe], have different tags (e.g. 1 and 2), then they have an "OR" relationship. In this case, node safe is reachable from [clo] if safe can be reached from either [rpa] or [na] which can be reach from [clo], i.e. node safe can be reached via either of the paths.

Referring to the notion of reachability, we can now define the set of goals satisfied by a decision as follows.

**Definition 3.7.** Given a DGC $CG = (N, E, C)$, $N = N_d \cup N_{int} \cup N_g$, in which $N_d$ and $N_g$ are the decisions and respective goals, a decision $d \in N_d$ meets (in the context of $C$) a goal $g \in N_g$ if $g$ is reachable from $d$ in the context $C$.

We use $\Gamma(d) = S, S \subseteq N_g$ to denote the set of goals met by decision $d$.

**Example 3.8.** (Example 3.4 continued.) As shown in Fig. 1, in the given context $C = \{\rightarrow gastritis, \rightarrow clo\_allergy, \rightarrow inhibitor\} \cup C_r$:
- $rpa$ is reachable from [clo] and [asp];
- $na$ is reachable from [asp];
- $se$ is reachable from [asp] and [dn];
- $safe$ is reachable from [rpa, na] and hence [asp];
- $cheap$ is reachable from $se$ and hence [asp] and [dn];
- $na$, and hence $safe$, is not reachable from [clo] since edge [clo|na] is inapplicable.

We can derive that $\Gamma(asp) = \{safe, cheap\}$ and $\Gamma(dn) = \{cheap\}$, i.e. decision $asp$ meets the goal $safe$ and $cheap$ while decision $dn$ only meets the goal $cheap$; and the decision $clo$ meets neither of the two goals.

According to the definitions given above, we can see that, given a DGC, whether a decision can meet a goal depends on the reachability of the goal from the decision. This reachability in turn depends on whether there exists paths leading from the decision to the goal and the applicability of the defeasible edges in the paths. Even for two DGCs with exactly the same acyclic graph, i.e. the same $N$ and $E$, a change in context $C$ may render some formerly applicable defeasible edges inapplicable or some inapplicable ones applicable, and hence may change the goals that can be met by the decisions. We modify Example 3.2 to illustrate how different contexts would affect the goals met by the decisions.

**Example 3.9.** Suppose the inhibitor is currently out-of-stock, i.e. ($\rightarrow \neg inhibitor$), while other contexts remain the same. In the new contexts, we now have $E_{ia} = \{[clo|na], [rpa|safe]\}$. Thus, the goal safe is no longer reachable from [rpa, na] and hence from [asp]. We can derive that $\Gamma(asp) = \{cheap\}$ and $\Gamma(dn) = \{cheap\}$. Both decision $asp$ and $dn$ meet the goal $cheap$ while no goal meets the goal safe.

Comparing Example 3.8 and 3.9, we can see that just by changing the contexts, the goals met by the decisions also change.

With the ability to capture context information, DGC offers greater expressiveness and flexibility in modelling decision problems. As an example, we show that DGCs generalise Decision Frameworks as follows.

**Proposition 3.10.** For any decision framework $DF = (D, A, G, T_{DA}, T_{DG})$, there is a DGC $CG = (N, E, C)$, $N = N_d \cup N_{int} \cup N_g$, with $N_d = D$, the decisions, $N_g = G$ the goals, such that for each $d \in D$, $g \in G$, it holds that $g \in \Gamma(d)$ in $DF$ if and only if $g \in \Gamma(d)$ in $CG$.

**Proof.** $g_k \in \Gamma(d_i)$ in $DF$ if and only if there exists an attribute $a_j \in A$, such that $T_{DA}[a_j] = 1$ and $T_{DG}[k, j] = 1$. $T_{DA}[a_j] = 1$ if and only if there exists an applicable edge $[d_i|a_j]$ in $CG$, and $T_{DG}[k, j] = 1$ if and only if there exists an applicable edge $[a_j|g_k]$ in $CG$. By Definition 3.5 and 3.7, $g_k$ is reachable from $d_i$. Hence, $d_i$ meets $g_k$ in $CG$.

This proposition holds as DGCs are capable of capturing any decision-attribute-goal relations that can be captured by DFs.

## 4 MAKING CONTEXT-BASED DECISIONS

After representing the decision problem as a DGC defined in the previous section, we illustrate how to map a DGC to an ABA framework and determine whether a decision is "good" with the ABA framework. An ABA framework can incorporate both the decision problem and the decision criteria simultaneously. The problem of identifying "good" decisions in a DGC can then be transformed into the problem of determining the admissibility of arguments in the corresponding ABA framework.

With the decisions meet goals information, different decision criteria can be adopted to evaluate the decisions. Dominance is used in [11]. In this paper, we focus on dominant decisions.
Following [11], a decision \(d \in N_d\) is dominant iff it meets all goals that are ever met by other decisions, i.e. let \(S = \Gamma(d)\), there exists no \(d'\) such that \(d' \neq d, g' \in \Gamma(d')\) and \(g' \notin N_S \setminus S\).

**Definition 4.1.** Given a DGC \(CG = (N, E, C)\), \(N = N_d \cup N_int \cup N_g\) with \(N_d\) the decisions, \(N_g\) the goals, \(E = E_S \cup E_d\) with \(E_S\) the strict edges, \(E_d\) the defeasible edges, \(C = C_p \cup C_t\) with \(C_p\) the context primitives, and \(C_t\) the context rules, the Dominant ABA Framework drawn from \(CG\) is \(ABF = (L, R, A, C)\), in which:

- \(R\) is such that:
  - for all \(e = [n|n'] \in E_d\): edge\((n, n', t(e)) \leftarrow \in R\);
  - edge\((x, y, t) \leftarrow dEdge(x, y, t) \in R\);
  - reach\((x, y) \leftarrow \neg \text{edge}\((x, y, t) \in R\);
  - reach\((x, y) \leftarrow \neg \text{reach}\((x, w_1), \text{edge}\((w_1, y, t)\),
    \neg \text{unreachableSib}\((w_1, y, t, x) \in R\);
  - \text{unreachableSib}\((w_1, y, t, x) \leftarrow \text{edge}\((w_2, y, t), \neg \text{reach}\((x, w_2)\), \neg \text{met}\((d, g), \text{notDom}(d, g), \text{othersMet}(d, g) \in R);\)
  - \text{notDom}(d, g) \leftarrow \neg \text{met}(d, g), \text{othersMet}(d, g) \in R;\)
  - \text{notOthers}(d, g) \leftarrow \neg \text{met}(d, g), \text{othersMet}(d, g) \in R;\)
  - \text{notMet}(d, g), \text{notOthers}(d, g), \text{notMet}(d_1, g), \cdots, \text{notMet}(d_n, g) \in R;\)
  - \text{for all} \(n \in N_t\) \(X \in \{(1) \text{ direct edge from } x \text{ to } y; (2) \text{ through intermediate node } w_1 \text{ and its sibling } w_2\).\)

The intuition of Definition 4.1 is the following. We know a decision \(d\) meets a goal \(g\) if node \(g\) is reachable from the set \(\{d\}\). We also know that a node \(y\) is reachable from a set of nodes \(x\) under either of the two conditions, as illustrated in Fig. 2:

1. if there exists an edge leading from node \(x\) to \(y\): reach\((x, y) \leftarrow \text{edge}\((x, y, k)\);
2. there is an intermediate node \(w_1\) such that:
   - \(w_1\) is reachable from \(x\), and there exists an edge from \(w_1\) to \(y\) that is also tagged with \(t\), and
   - if \(w_1\) has a "sibling" node \(w_2\), there is an edge from \(w_2\) to \(y\) that is also tagged with \(t\), then: \(w_2\) is also reachable from \(x\), given by rules:
      - reach\((x, y) \leftarrow \text{reach}\((x, w_1), \text{edge}\((w_1, y, t)\),
        \neg \text{unreachableSib}\((w_1, y, t, x) \leftarrow \text{edge}\((w_2, y, t), \neg \text{reach}\((x, w_2)\).\)

A decision \(d\) is dominant if it meets all goals that are ever met by other decisions. Hence, the two premises of \(\text{notDom}(d)\) are

\(\text{notMet}(d, g)\) and \(\text{othersMet}(d, g)\), representing "the decision does not meet this goal" and "some other decisions can meet this goal" respectively. The contrary of "some other decisions can meet this goal" is "no other decisions meet this goal", represented by the rule:

\(\text{notOthers}(d, g) \leftarrow \text{notMet}(d, g), \cdots, \text{notMet}(d_n, g)\).

Each defeasible edge \(e = [n|n']\) has an associated assumption \(\neg \text{edge}\((n, n', t(e))\) with a contrary \(\text{edge}\((n, n', t(e))\) for some \(x = n, y = n'\) and \(t = t(e)\). Decision contexts in \(C\) are represented as elements of \(R\) in the ABA framework. With such formalization, context information can be captured by the framework and influence the decision making by moderating the applicability of the defeasible edges. For example, if a defeasible edge \(e = [n|n']\) tagged with \(t(e) = i\) becomes inapplicable in a new context \(C'\), then we have \(\{\} \leftarrow \neg \text{edge}\((n, n', i)\) which can form an attack on the assumption \(\text{edge}\((n, n', i)\).

We already know that the problem of identifying "good" decisions in a decision graph is equivalent to the problem of determining the admissibility of the decisions in the corresponding ABA framework [21]. This also applies to a DGC, which is formally proved as follows:

**Proposition 4.2.** Given a DGC \(CG = (N, E, C)\), \(N = N_d \cup N_int \cup N_g\) with \(N_d\) the decisions, \(N_g\) the goals, let \(ABF\) be the dominant ABA framework drawn from \(CG\). Then for all decisions \(d \in N_d\), \(d\) is dominant in \(CG\) if \(\{\text{dom}(d)\} \leftarrow \text{dom}(d)\) is admissible in \(ABF\).

**Proof.** (Sketch.) First, we prove dominance implies admissibility for \(d_i \in N_d\). Since \(d_i\) is dominant, for each goal \(g_j\), either (1) \(d_i\) meets goal \(g_j\), therefore argument \(\{\} \leftarrow \text{met}(d_i, g_j)\) exists and is not attacked; or (2) there is no argument \(\{\} \leftarrow \text{met}(d_k, g_j)\) for all \(d_k \in N_d\), therefore argument \(\{\} \leftarrow \text{notOthers}(d_i, g_j)\) exists and is not attacked. In both cases, the attackers of the argument \(\{\} \leftarrow \text{dom}(d_i)\) are always counter attacked. Thus, \(\{\} \leftarrow \text{dom}(d_i)\) withstands all attacks. Moreover, since \(\{\} \leftarrow \text{notOthers}(d_i, g_j)\) is conflict-free, \(\{\} \leftarrow \text{dom}(d_i)\) is admissible.

We then prove admissibility implies dominance. Since the argument \(\{\} \leftarrow \text{dom}(d_i)\) is admissible, all its attackers, i.e. \(\{\text{notMet}(d_i, g_j), \text{othersMet}(d_i, g_j)\} \leftarrow \text{notDom}(d_i)\) for all \(g_j \in N_g\), must be counter attacked. We know that each such attacker is counter attacked either because there exists an argument \(\{\} \leftarrow \text{met}(d_i, g_j)\) or argument \(\{\} \leftarrow \text{notOthers}(d_i, g_j)\), i.e. either \(d_i\) meets \(g_j\) or no other \(d_k \in N_d\) meets \(g_j\). Thus, \(d_i\) is dominant. □
Proposition 4.2 shows that ABA frameworks can be used to identify dominant decisions.

Example 4.3. (Example 3.8 continued.) Given the DGC in Fig. 1, a dominant ABA framework \((\mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C})\) can be constructed according to Definition 4.1, with the variables \(x, y, w_1, w_2\) instantiated to the elements in \(N\), variable \(t\) instantiated to the tags, variable \(d\) instantiated to the elements of \(N_d\) and variable \(g\) instantiated to the elements of \(N_g\) in the DGC. Due to space limitation, we list a few elements of each component instead of the fully instantiated framework here:

\[R\] consists of:
- \(\text{edge}(\text{clo}, \text{rpa}, 1) \leftarrow \text{edge}(\text{asp}, \text{na}, 2) \leftarrow \text{edge}(\text{dn}, \text{se}, 2) \leftarrow \text{edge}(\text{rpa}, \text{safe}, 1) \leftarrow \cdots\)
- \(\text{edge}(\text{clo}, \text{na}, 1) \leftarrow \text{dEdge}(\text{clo}, \text{na}, 1)\)
- \(\text{edge}(\text{asp}, \text{na}, 2) \leftarrow \text{dEdge}(\text{asp}, \text{na}, 2)\)
- \(\text{edge}(\text{rpa}, \text{safe}, 1) \leftarrow \text{dEdge}(\text{rpa}, \text{safe}, 1) \cdots\)
- \(\text{reach}(\text{clo}, \text{rpa}) \leftarrow \text{edge}(\text{clo}, \text{rpa}, 1) \cdots\)
- \(\text{reach}(\text{clo}, \text{safe}) \leftarrow \text{reach}(\text{clo}, \text{rpa}), \text{edge}(\text{rpa}, \text{safe}, 1) \cdots\)
- \(\neg\text{unreachableSib}(\text{rpm}, \text{safe}, 1, \text{clo}) \cdots\)
- \(\neg\text{unreachableSib}(\text{rpm}, \text{safe}, 1, \text{clo}) \leftarrow \text{edge}(\text{na}, \text{safe}, 1) \neg\text{reach}(\text{clo}, \text{na}) \cdots\)
- \(\text{notDom}(\text{asp}) \leftarrow \text{notMet}(\text{rpm}, \text{safe}), \text{othersMet}(\text{rpm}, \text{safe}) \cdots\)
- \(\text{noOthers}(\text{rpm}, \text{safe}) \leftarrow \text{notMet}(\text{clo}, \text{safe}), \text{notMet}(\text{dn}, \text{safe}) \cdots\)
- \(\neg\text{dEdge}(\text{rpm}, \text{safe}, 1) \leftarrow \text{gastritis}, \neg\text{inhibitor}\)
- \(\text{clo_allergy} \leftarrow \neg\text{dEdge}(\text{clo}, \text{na}, 1) \leftarrow \text{clo_allergy}\)
- \(\text{gastritis} \leftarrow \neg\text{dEdge}(\text{asp}, \text{na}, 2) \leftarrow \text{asp_allergy}\)
- \(\text{inhibitor} \leftarrow \cdots\)

\[A\] consists of:
- \(\neg\text{unreachable}(\text{rpm}, \text{safe}, 2, \text{clo}) \neg\text{unreachable}(\text{rpm}, \text{safe}, 2, \text{clo}) \cdots\)
- \(\neg\text{reach}(\text{clo}, \text{rpm}) \neg\text{reach}(\text{clo}, \text{rpm}) \neg\text{reach}(\text{clo}, \text{safe}) \neg\text{reach}(\text{clo}, \text{safe}) \cdots\)
- \(\neg\text{reach}(\text{rpm}, \text{safe}) \neg\text{reach}(\text{rpm}, \text{safe}) \neg\text{reach}(\text{rpm}, \text{safe}) \cdots\)
- \(\text{dom}(\text{clo}) \neg\text{dom}(\text{asp}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \cdots\)
- \(\text{notMet}(\text{rpm}, \text{safe}) \text{notMet}(\text{rpm}, \text{safe}) \text{notMet}(\text{rpm}, \text{safe}) \cdots\)
- \(\text{othersMet}(\text{rpm}, \text{safe}) \text{othersMet}(\text{rpm}, \text{safe}) \text{othersMet}(\text{rpm}, \text{safe}) \cdots\)
- \(\text{dEdge}(\text{clo}, \text{na}, 1) \text{dEdge}(\text{clo}, \text{na}, 1) \cdots\)
- \(\text{dom}(\text{clo}) \text{dom}(\text{asp}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \text{dom}(\text{dn}) \cdots\)

\[C\] is as defined in Definition 4.1.

In this ABA framework, argument \((\text{dom}(\text{asp})) + \text{dom}(\text{asp})\) is admissible. However, argument \((\text{dom}(\text{clo})) + \text{dom}(\text{clo})\) and argument \((\text{dom}(\text{dn})) + \text{dom}(\text{dn})\) are not admissible. By Proposition 4.2, decision \text{asp} is the only dominant decision.

Given a DGC \(CG = (N, E, C)\), a change in context has no impact on the graph part \((N, E)\), but affects the elements in the set \(C\), which are represented as rules in \(R\) in the corresponding ABA framework. Hence, to obtain a new ABA framework which corresponds to the new context, we only need to update the elements in \(R\), by replacing the old rules affected with new ones derived from the new context. As we will see in the following example, as context changes, the applicability of defeasible edges may change which may affect the reachability of goal nodes from decision nodes. Thus, the dominance of decisions may be different when context varies.

Example 4.4. (Example 3.9 continued.) When the inhibitor is out-of-stock, the dominant ABA framework that can be drawn from the DGC in this example is almost the same \((\mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C})\) as given in Example 4.3, but with the following modification:

- Remove the rule \(\text{inhibitor} \leftarrow \text{inR}\);
- Add the rule \(\neg\text{inhibitor} \leftarrow \text{inR}\).

In this context, both \((\text{dom}(\text{asp})) + \text{dom}(\text{asp})\) and \((\text{dom}(\text{dn})) + \text{dom}(\text{dn})\) are admissible. Hence, by Proposition 4.2, decisions \text{asp} and \text{dn} are dominant.

From Example 4.3 and 4.4, we can see that when the context changes, the dominance of the decisions may also change as a result. Thus, context can affect the decision. In other words, different decisions may be chosen in different contexts.

5 EXPLAINING NON-DOMINANT DECISIONS

The process of making decisions with argumentation paves the way for generating meaningful explanations for the decisions. For each non-dominant decision, it is useful to identify the reasons why it is not a “good” decision in the given context. In this section, we describe two types of explanations, argument explanation (arg-explanation) and context explanation (cont-explanation) for non-dominant decisions. Arg-explanations focus on identifying the cause of non-admissibility. By Proposition 4.2, we know that choosing dominant decisions in a DGC is equivalent to identifying admissible arguments in the corresponding ABA framework. In other words, if a decision is non-admissible, it will not be selected. Thus, identifying the cause of non-admissibility can help to explain why a decision is not selected. Cont-explanations provide more informative explanations when the non-admissibility can be traced to the contexts, by telling whether the failure of a decision can be attributed to the contexts and why it fails in this context.

Based on the ABA framework drawn from a DGC, as defined in Definition 4.1, we formalize the arg-explanations for non-admissible ABA arguments in Definition 5.1. Note that, the following definition is building on the explanations for AA arguments [15] given in the background, and the definition of the pruning operator, \(\setminus\), also follows that given before in the Background.

Definition 5.1. Given an ABA framework \(ABF = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C})\) with the corresponding AA framework \(AF = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \mathcal{C})\), let \(a \in B\) be a non-admissible argument in \(AF\). Then, \(B \subseteq B'\) is an explanation of \(a\) if and only if the following two conditions hold:

1. \(a\) is admissible in \(AF \setminus B\), and
2. there exists no \(B'' \subset B\) such that \(a\) is admissible in \(AF \setminus B''\). If no such \(B\) exists, \((a)\) is the arg-explanation of \(a\).

The non-admissibility of an argument \(a\) can be attributed to a set of attackers \(B\) from which \(a\) cannot be defended. The intuitive idea of Definition 5.1 is: if we remove all arguments in set \(B\) from the framework and argument \(a\) becomes admissible in the pruned framework, then \(B\) contains all attackers from which \(a\) cannot be defended and they constitute an arg-explanation for why \(a\) is not admissible in the unpruned framework. Note that set \(B\) needs to be minimal. In the cases when no such set \(B\) exists, the non-admissibility of \(a\) can only be attributed to itself, i.e. when \(a\) attacks itself.

Although arg-explanations are computed using admissibility semantics, they can provide interpretable reasons to explain the
decisions. We continue to use the dominant decision criterion as an illustration. For a non-dominant decision $d$, which fails to meet all goals that are ever met by other decisions, arguments in an arg-explanation of $\{dom(d)\}$ identify goals that are not met by $d$ but met by some other decisions. Formally:

**Proposition 5.2.** Given a DGC $CG = (N, E, C)$, $N = N_d \cup N_g \cup N_{int}$, with $N_d$ the decisions, $N_g$ the goals, let $ABF = \langle L, R, A, C \rangle$ be the dominant ABA framework drawn from $CG$. Then, for all $R_a = \{dom(d)\} \vdash dom(d)$ in ABF where $d \in N_d$ is not dominant,

1. if for some $g \in N_g$, $\{notMet(d, g), othersMet(d, g)\} \vdash notDom(d)$ is in an arg-explanation of $R_a$, then $g$ is not met by $d$;
2. if for some $g \in N_g$, argument $\{\} \vdash \neg \text{Edge}(n_i, n_j, 1)$ is in an arg-explanation of $R_a$ and is attacking an argument $\Delta_1 \vdash \text{met}(d, g)$, then $g$ is not met by $d$;
3. if for some $d' \in D$, $d' \neq d$ and $g \in N_g$, $\Delta_2 \vdash \text{met}(d', g)$ is in an arg-explanation $R_a$, then $d$ does not meet $g$ and $d'$ meets $g$.

**Proof.** By the construction of ABF (Definition 4.1), for a decision $d$ and a goal $g$, argument $A = \{notMet(d, g), othersMet(d, g)\} \vdash notDom(d)$ exists in ABF. In order to defend the root argument $R_a = \{dom(d)\} \vdash dom(d)$, argument $A$ needs to be counter-attacked by either argument $C = \Delta_1 \vdash \text{met}(d, g)$ or $E = \Delta_3 \vdash notOthers(d, g)$. For (1), it is easy to see that if $A$ is removed, all its sub-level attackers and defenders are also removed, and thus $R_a$ would become admissible. Thus, $A$ is an arg-explanation of $R_a$ and $d$ does not meet $g$. To show (2), we observe that since $\neg \text{Edge}(n_i, n_j, 1)$ attacks argument $C$, $d$ does not meet $g$. Similarly for (3), since the argument attacked by $F = \Delta_2 \vdash \text{met}(d', g)$ is $\Delta_3 \vdash notOthers(d, g)$ which in turn attacks argument $A$, $F$ defends $A$ and $d$ does not meet $g$. Also, since argument $F$ exists in ABF, $d'$ meets $g$.

In order for a decision $d$ to be dominant, for each goal $g$, it must be either $d$ meets $g$ or all decisions do not meet $g$. Thus, it is easy to see that $d$ does not meet $g$ and some other decision meets $g$ would be in an arg-explanation. The inapplicability of a defeasible edge can also be in an arg-explanation as it may affect the reachability of $g$ from $d$.

**Example 5.3.** (Example 4.3 continued.) Given the ABA framework $ABF$ in Example 4.3, the corresponding AA framework constructed from $ABF$ for the decision administer clopidogrel ($clo$) is shown in Fig.3. There are six arg-explanations $\{A, B\}$, $\{A, H\}$, $\{B, D\}$, $\{B, F\}$, $\{D, H\}$ and $\{F, H\}$ for the root argument $R_a$. Removing any one of them from this AA framework makes the argument $R_a$ admissible. By Proposition 5.2, we interpret this arg-explanation as follows:

The decision of administering clopidogrel is not dominant as it does not meet the goal safe or the goal cheap.

An arg-explanation for a decision contains the reasons for the non-admissibility of the argument that embeds the decision. In other words, it contains the reasons why the decision is not dominant, which may or may not be due to contexts in which the decision is made. Hence, it is useful to know why the decision is non-admissible in a particular context. We use cont-explanation, which only contains the reasons that can be traced down to the decision contexts, for this purpose. Formally:

**Definition 4.5.** Given a DGC $CG = (N, E, C)$, $N = N_d \cup N_g \cup N_{int}$, with $N_d$ the decisions, $N_g$ the goals, let $ABF = \langle L, R, A, C \rangle$ be the dominant ABA framework drawn from $CG$. Let $d \in N_d$ be a non-dominant decision in $CG$ and the set of arguments $E_{arg}$ be an arg-explanation for the non-admissible argument $R_a = \{dom(d)\} \vdash dom(d)$ in ABF. Then for each argument $\{\} \vdash \neg \text{Edge}(n_i, n_j, 1)$ in $E_{arg}$, $R \subseteq R$ is in a cont-explanation $E_{con}$ of $R_a$.

According to Definition 5.4, if an argument has a cont-explanation, it must also have an arg-explanation, but not vice versa. A cont-explanation is derived from the ABA rules in arguments of inapplicable edges. It is worth noting that inapplicable edges are not always relevant to the computation of cont-explanations. Only the ones that affect the reachability of goals from the concerned decision are relevant to the derivation of cont-explanations. Formally:

**Proposition 5.5.** Given a DGC $CG = (N, E, C)$, $N = N_d \cup N_g \cup N_{int}$, with $N_d$ the decisions, $N_g$ the goals, let $ABF = \langle L, R, A, C \rangle$ be the dominant ABA framework drawn from $CG$, in which

- $d \in N_d$ is a non-dominant decision in $ABF$ that fails to meet some goal $g \in N_g$, and there exists an argument $\Delta \vdash \text{met}(d, g)$;
- $d' \in N_d$ meets $g$, $d' \neq d$;
- $\text{edge}(n_i, n_j) \in E_{cl}$ is inapplicable in the context $C$ and the argument for its inapplicability is $\{\} \vdash \neg \text{Edge}(n_i, n_j, t)$, $R \subseteq R$.

Then $R$ is in a cont-explanation of $\{dom(d)\} \vdash dom(d)$ if and only if $\text{Edge}(n_i, n_j, t) \in \Delta$.

**Proof.** An intuitive proof of Proposition 5.5 is as follows. According to Definition 5.4, if $R$ is in a cont-explanation of $R_a = \{dom(d)\} \vdash dom(d)$, argument $E = \{\} \vdash \neg \text{Edge}(n_i, n_j, t)$ must be in an arg-explanation of $R_a$. Thus, $E$ must be an attack on the argument $D = \Delta \vdash \text{met}(d, g)$, and $D$ cannot be defended from $E$. Since in ABA attacks can only be directed at the assumptions in the support of an argument, $\text{Edge}(n_i, n_j, t)$ must be in the support $\Delta$ of $D$. Conversely, if $\text{Edge}(n_i, n_j, t)$ is in the support of argument $D$, then...
then argument $E$ forms an attack from which $D$ cannot be defended and renders $R_u$ non-admissible. Thus, $E$ is in an arg-explanation of $R_u$ and $R$ is in a cont-explanation of $R_u$.

**Example 5.6.** (Example 5.3 continued.) The argument $\{ \} \vdash_R \neg \text{dEdge}\{\text{clo}, \text{na}, 1\}$ where $R = (\neg \text{dEdge}\{\text{clo}, \text{na}, 1\} \leftarrow \text{clo\_allergy}, \text{clo\_allergy} \leftarrow \text{dEdge}\{\text{clo}, \text{na}, 1\})$ is in the arg-explanation of $R_u = \{\text{dom}\{\text{clo}\}\} \vdash \text{dom}\{\text{clo}\}$. By Definition 5.4, the cont-explanation for $R_u$ is $E_{\text{cont}} = (\neg \text{dEdge}\{\text{clo}, \text{na}, 1\} \leftarrow \text{clo\_allergy}, \text{clo\_allergy} \leftarrow)$. According to Proposition 5.2, we can interpret this cont-explanation as follows:

*The decision of administering clopidogrel does not meet the goal safe as the patient has an allergy to clopidogrel.*

We can also observe that in this example, $\text{dEdge}\{\text{clo}, \text{na}, 1\}$ is in the support of argument $\Delta \vdash \text{met}\{\text{clo}, \text{safe}\}$.

### 6 RELATED WORK

Researches on context-based decision making have focused on building knowledge representation models and developing decision making methods. Logical models, such as ontologies [6, 30] and logic rules [19, 22, 27], have been widely used due to their high interpretability and expressibility. Subsumption checking in an ontology is used to perform activity recognition in [6]. A first order logic model is introduced in [27] to express complex rules with context. Several Bayesian approaches [18, 25] have also been proposed to model the decision making process. A combined approach is adopted in [5], which employs ontologies and logic rules for representation and makes decisions with Markov Logic Networks.

In [20] and [31], case-based reasoning is used to generate context awareness by referring back to similar previous scenarios. Different from these works, we take an argumentation-based approach, which enjoys the benefits of transparent decision making and ease of explanation generation.

In an argumentation-based approach, argumentation has been used as a formalism for representing decision problems as well as a reasoning mechanism for computing decisions guided by some decision criteria. In [1], arguments in favour or against each decision alternative are constructed, and then evaluated against a pessimistic or an optimistic criterion. In [2], arguments for decisions are built using AA and their acceptability are evaluated with the classical or an optimistic criterion. Similar to our approach, a unified approach is adopted in [5], which employs ontologies and logic rules for representation and makes decisions with Markov Logic Networks.

As our proposed approach is more for the case of a single-agent system, we are interested to explore how it can be applied to a multi-agent system. ABA dialogues and frameworks can be used to communicate arguments and compute decisions in a multi-agent system. In [16], the agents exchange arguments in the form of an ABA dialogue. A joint ABA framework can be constructed from the dialogue to model the decision-making process of two agents. Our proposed approach is also based on ABA framework. Hence, theoretically, the methods we used to make context-based decisions and to generate the two types of explanations can also be applied to the joint ABA framework for two or more agents. However, protocols are required to govern the dialogues among agents. Also, since agents may have different goals and candidate decisions, multiple Decision Graphs with Context (DGC) (one for each agent) may be needed to model the problem.

We have only focused on explanations for non-admissible decisions in this paper. Currently, we are working on generating the two types of explanations for admissible decisions. Some of our discussions regarding the applicability of edges share similar ideas as the reasoning in higher order argumentation semantics, such as [3, 17, 24, 26], where arguments are allowed to attack relations and other arguments. In the future, we will study the relevancy between our approach and these higher order argumentation semantics. We are also interested in incorporating preferences in addition to contexts.
ACKNOWLEDGMENTS

This research is supported by the National Research Foundation, Prime Minister’s Office, Singapore under its IDM Futures Funding Initiative, the Singapore Ministry of Health under its National Innovation Challenge on Active and Confident Ageing (NIC Project No. MOH/NIC/COG04/2017), the Interdisciplinary Graduate School, Nanyang Technological University under its Graduate Research Scholarship and the Alibaba-NTU Singapore Joint Research Institute.

REFERENCES