Resolving Conflicts in Clinical Guidelines using Argumentation*

Kristijonas Čyras
Imperial College London
London, United Kingdom
k.cyras@imperial.ac.uk

Tiago Oliveira
National Institute of Informatics
Tokyo, Japan
toliveira@nii.ac.jp

ABSTRACT

Automatically reasoning with conflicting generic clinical guidelines is a burning issue in patient-centric medical reasoning where patient-specific conditions and goals need to be taken into account. It is even more challenging in the presence of preferences such as patient's wishes and clinician's priorities over goals. We advance a structured argumentation formalism for reasoning with conflicting clinical guidelines, patient-specific information and preferences. Our formalism integrates assumption-based reasoning and goal-driven selection among reasoning outcomes. Specifically, we assume applicability of guideline recommendations concerning the generic goal of patient well-being, resolve conflicts among recommendations using patient's conditions and preferences, and then consider prioritised patient-centered goals to yield non-conflicting, goal-maximising and preference-respecting recommendations. We rely on the state-of-the-art Transition-based Medical Recommendation model for representing guideline recommendations and augment it with context given by the patient's conditions, goals, as well as preferences over recommendations and goals. We establish desirable properties of our approach in terms of sensitivity to recommendation conflicts and patient context.

KEYWORDS

Medical reasoning; Structured argumentation; Ariadne principles

ACM Reference Format:

Kristijonas Čyras and Tiago Oliveira. 2019. Resolving Conflicts in Clinical Guidelines using Argumentation. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 9 pages.

1 INTRODUCTION

Medical reasoning involves careful deliberation about the condition of a patient and possible treatments. Clinical guidelines provide best practice recommendations for achieving patient well-being given a disease and describe management of a generic patient, recommending multiple options to choose among, given a concrete patient and their context. When managing multiple health conditions (multimorbidities), guidelines need to be merged, whence multiple interactions must be considered, as they influence the evolution of the patient [12, 13]. In particular, the recommended actions may be inapplicable, conflicting, overlapping, and so forth. It is hard for clinicians to follow the best practices in the presence of conflicts among guidelines. In such settings, knowledge representation methods come handy, particularly in representation of guidelines and their interactions.

Transition-based Medical Recommendation model (TMR) [42] is a state-of-the-art [31] development in representation of clinical guideline recommendations. TMR identifies components and relations typically present in multimorbidity situations, such as clinical care actions and their effects on physical properties. To capture interactions when merging guidelines, in [42] a mechanism is advanced for identifying relationships among multiple recommendations, such as contradiction, repetition, alternative. TMR thus offers a comprehensive template for clinical guidelines and their interactions. Yet, TMR does not afford a method for representing patient-specific information. More importantly, TMR does not afford *reasoning* mechanisms to determine which recommendations to follow for a given patient.

In general, whereas representation of clinical guidelines is a well-advanced area, automated reasoning with those representations, especially in the presence of conflicts, is a limiting factor [12, 28, 31]. An additional hurdle is taking into account the context of the patient, pertaining to patient-specific conditions, patient-centric goals and preferences from various parties involved [28, 32, 38]. For instance, Ariadne principles [25] show the importance and difficulty of integrating interaction assessment, individual management and patient's and/or clinician's preferences in multimorbidity setting. We here advance a framework to address the above mentioned issues by applying an argumentation-based method to allow an autonomous agent to reason with conflicting clinical guidelines in the context of patient information, goals and various preferences.

Generally speaking, argumentation allows to reason with incomplete and conflicting information in a way that aims to emulate human reasoning. Argumentation is used for modelling reasoning of autonomous agents in multi-agent systems, see e.g. [5, 17, 27, 30], and has been extensively applied in the medical domain, see e.g. [11, 15, 21, 26, 35, 37]. In medical reasoning particularly, "argumentation is appealing as it allows for important conflicts to be highlighted and analysed and unimportant conflicts to be suppressed." [3] We propose to use the structured argumentation (see e.g. [6]) formalism *Assumption-Based Argumentation with Preferences* (ABA^+) [7, 8] for automating patient-centric reasoning with conflicting guideline recommendations and preferences. This choice is motivated by the simplicity of knowledge representation in ABA^+ and its dealing with preferences differently than other formalisms: it reverses attacks due to preferences and, importantly, in doing so preserves conflict-freeness of sets of assumptions; these aspects allow for a leaner representation while yielding desirable properties, as demonstrated further ahead. ABA^+ also has known complexity results [9, 19] and a working implementation [4].

*The authors have made equal contributions and are listed alphabetically.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

ABA^+ is deployed to use TMR for representation of recommendations and interactions via rules and arguable elements from which arguments (as deductions) are constructed. ABA^+ augments this representation with patient-specific information and uses extension-based argumentation semantics for reasoning. It thus provides an assumption-driven reasoning method whereby assumed applicability of recommendations is argued about using patient's conditions. At the same time, ABA^+ deals with preferences over actions (equivalently, recommendations) as specified by e.g. the patient. We establish that the resulting recommendations (corresponding to extensions of ABA^+ frameworks) are non-conflicting. We also augment ABA^+ to form ABA^+G, incorporating a goal-driven reasoning mechanism to determine the best non-conflicting recommendations given patient-centric goals ordered by importance. This allows our approach to meet Ariadne principles. We also illustrate our approach with a case study from [42] and obtain arguably desirable outcomes when reasoning with conflicting recommendations in the context of a patient.

The paper is structured thus. In Section 2 we consider desiderata for our approach in terms of patient management principles from medical literature. We then in Section 3 describe the problem of reasoning with interacting recommendations in the context of a patient. In Section 4 we propose to use ABA^+ and its development ABA^+G for assumption-based patient-centric reasoning with recommendations, goals and preferences. We discuss related work in Section 5 and finish with conclusions and future work in Section 6.

2 PRINCIPLES OF PATIENT MANAGEMENT

We situate our work in the context of principles of patient management in multimorbidity setting. Several works mention various principles for patient management [12, 13, 28, 32, 38], but do so in a loose and fragmented manner. Hence, it is difficult to acquire a comprehensive understanding of what these principles should be. [25] is among the very few works with a comprehensive enumeration and description of such principles, called **Ariadne principles**. Our interpretation of them is as follows.

1. **Interaction assessment:** recommendation interactions and respective effects are identified and resolved. In contrast to patients with a single disease, when managing patients with multimorbidities, a variety of potential interactions between diseases and treatments may occur and worsen the course of the disease(s).
2. **Prioritisation and patient preferences:** to guide the reasoning, priorities among goals are established while respecting the patient's preferences and state. These priorities and preferences are used to consolidate heavy treatment burdens and competing treatment goals. Treatment goals are expressed in terms of symptom relief, disease prevention, avoidance of undesired outcomes, and preservation or improvement of life expectancy and quality.
3. **Individualised management:** a treatment plan as a set of recommendations is devised in accordance with the patient's state, preferences and the prioritised goals. This plan should provide non-conflicting recommendations for the given patient.

Ariadne principles do not provide specific methods to solve recommendation conflicts or to elicit preferences, but rather state which dimensions should be accounted for while reasoning. Regarding treatment goals, information about the impact of treatments on life expectancy and quality of life may not (or is often not) available. Therefore, limiting treatment goals to symptom relief, disease prevention, avoidance of undesired outcomes—i.e., effects brought about (or not) by treatments—seems to be the most practical choice in these circumstances. Furthermore, the patient's preferences over actions and the clinician's priorities over goals should result from a discussion between the patient and the physician, and need to be taken into account when devising a treament plan for the patient.

For an autonomous agent to reason with conflicting medical recommendation within the TMR model, we need to establishing foundations for that reasoning in a setting of patient management. TMR provides an expressive representation template for clinical guideline recommendations in discordant multimorbidity, respective interaction types, and several measures such as causation belief, deontic strength, and evidence level. However, it does not demonstrate the aggregation of these elements in reasoning to produce treatment solutions for specific patients. Devising such solutions is no trivial task not only due to the complexity brought about by the number of existing health conditions and recommendations but also due to the overall under-specification of how decisions should be made in this setting. In this work, we aim to meet Ariadne principles by adequately handling conflicting guideline recommendations afforded by the TMR model, while taking into account the context that includes patient-specific conditions, goals and preferences.

3 PROBLEM SETTING

We here situate the problem of reasoning with interacting clinical guideline recommendations in the context of a patient. We first review the Transition-based Medical Recommendation (TMR) model and interactions among recommendations. We then discuss the context of a patient.

3.1 TMR Model

We first give the TMR model together with guideline recommendation interaction representation. They will be used to construct ABA^+ frameworks for reasoning with guidelines. (As in [42], we assume that a set of guidelines is merged into a single guideline so that recommendations are delivered by the same larger guideline.)

3.1.1 Recommendations. Figure 1 depicts an instance of a graphical schema for representing recommendations in TMR. (Recommendation concerning NSAID is taken from a diabetes guideline, and recommendation concerning Aspirin is taken from an osteoarthritis guideline.) It consists of the following components.[1]

(i) Name, e.g. R_1, R_2, at the top of a rounded box.
(We write R_k instead of Rk.)
Henceforth, we refer to a recommendation by its name.

(ii) A unique associated **action** A, e.g. *Adm. Aspirin*, *Adm. NSAID* (where *Adm.* stands for Administer).

(iii) Deontic strength, which we denote by δ, is indicated by a thick labelled arrow and "reflects a degree of obligatoriness expected for that recommendation" [42, p. 82]. It takes values

[1]The formal description of recommendations, with components as functions/relations, is long, cumbersome, and unnecessary for the purposes of this paper. Instead, we give an intermediate representation, following the alternative formal description (and visualisation) in [42] of TMR instances, faithful to the original but omitting certain aspects (as indicated below), which carries the necessary aspects required in this work.

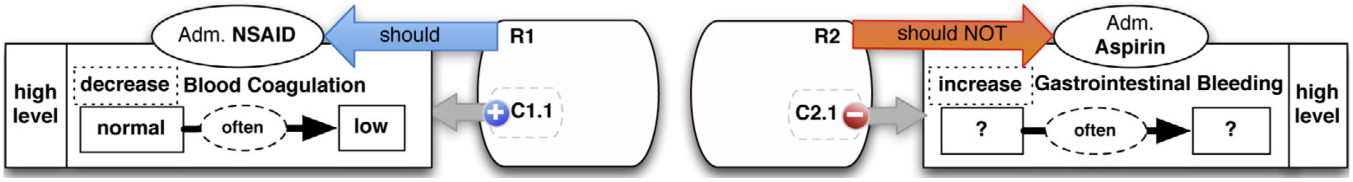

Figure 1: TMR representation schema instantiated with recommendations R_1 and R_2 [42, p. 83, Figure 2].

in $[-1, 1]$: if $\delta \geqslant 0$, then R recommends to perform the action; if $\delta < 0$, then R recommends to avoid the action. To discretise δ, the qualitative landmarks *must*, *should*, *may*, *should not*, *must not* corresponding to $1, 0.5, 0, -0.5, -1$, respectively, are used. E.g., the deontic strengths of R_1 and R_2 in Figure 1 are $\delta_1 = 0.5 =$ *should* and $\delta_2 = -0.5 =$ *should not*, respectively.

(iv) **Properties** that the associated action affects, e.g. *Blood Coag.*, *Gastro. Bleeding*. (If clear from the context, we abbreviate words as follows: e.g. *Coag.* and *Gastro.* abbreviate *Coagulation* and *Gastrointestinal*, respectively.)
In general, an action can affect more than one property P.

(v) **Effects** of the actions, e.g. *decrease*, *increase*.
An action A has one effect E on the property P it affects.

(vi) **Initial** and **final** values of the property that an action affects. For instance, *Adm. NSAID* leads to a *decrease* in *Blood Coag.* from the initial value *normal* to the final value *low*. Otherwise, ? represents *indeterminate* value.
In this paper we *will not* make use of, but mention for completeness, two quantitative values associated with an effect: *causation probability* – e.g. *often* – representing the likelihood of the action bringing the effect about; and *belief strength* – e.g. *normal level* – representing the level of evidence regarding bringing the effect about.

(vii) **Contributions** of the recommendation to the overall goals in the context of a guideline, e.g. $+C1.1$, $-C2.1$, indicated below the recommendation name.
A recommendation can have multiple contributions, each carrying an *identifier*, e.g. $C1.1$, $C2.1$, and *valued* in $[-1, 1]$ (indicating importance of achieving/avoiding the corresponding effect), discretised with signs $+$, $-$ and no sign, representing values greater than, less than and equal to 0, respectively.

An instance of TMR concerns a generic patient. In order to apply recommendations, one needs to consider specific patient *conditions*, pertaining to properties and the initial values of the effects that actions have on properties. For instance, a patient can have conditions *normal Blood Coag.* or *Gastro. Bleeding*. When using argumentation frameworks to reason with guidelines in Section 4, patient conditions will come as information additional to TMR instances. With the following intermediate representation of TMR instances we ensure that recommendations as well as patient-specific conditions will be representable in argumentation frameworks.

Definition 3.1. A **recommendation** is a tuple $(R, A, \delta, \mathcal{P}, \mathcal{E}, \mathcal{V}, \mathcal{C})$ consisting of the following components:

(i) name R,
(ii) action A,
(iii) deontic strength δ,
(iv) properties $\mathcal{P} = \langle P^1, \dots, P^n \rangle$ affected, for $n \geqslant 1$,
(v) effects $\mathcal{E} = \langle E^1, \dots, E^n \rangle$,
(vi) initial values $\mathcal{V} = \langle v^1, \dots, v^n \rangle$ of effects on properties,
(vii) contribution values $\mathcal{C} = \langle c^1, \dots, c^n \rangle$.

We identify any recommendation with its name R and with an abuse of notation may write $R = (R, A, \delta, \mathcal{P}, \mathcal{E}, \mathcal{V}, \mathcal{C})$. We use $\mathbb{R}$ to denote a fixed but otherwise arbitrary set of recommendations, unless specified otherwise.

Example 3.2. Recommendations $R_1 = (R_1, \textit{Adm. NSAID}, \textit{should}, \langle \textit{Blood Coag.} \rangle, \langle \textit{decrease} \rangle, \langle \textit{normal} \rangle, \langle + \rangle)$ and $R_2 = (R_2, \textit{Adm. Aspirin}, \textit{should not}, \langle \textit{Gastro. Bleeding} \rangle, \langle \textit{increase} \rangle, \langle ? \rangle, \langle - \rangle)$ are illustrated in Figure 1. So $\mathbb{R} = \{R_1, R_2\}$.

3.1.2 Interactions. Using TMR, Zamborlini et al. identify *interactions* among recommendations. Intuitively, interactions record the relationships between different recommendations, for instance, a *contradiction* relationship in case one recommendation urges to avoid the action suggested by another recommendation. Several types of interactions, such as contradiction, repetition, alternative, can be identified. We focus on the contradiction interaction in this paper, because it relates recommendations in direct conflict that can be naturally resolved by means of argumentation.

Contradiction interactions can be represented as triples (R, R', μ) with recommendations R and R', and the interaction's *modal strength* μ, which reflects the conclusiveness of the interaction. The interaction's modal strength can take two values, denoted by $\Box$ and $\Diamond$, where $\Box$ means 'the interaction will certainly occur if the related recommendations are prescribed' [42] and $\Diamond$ means 'the interaction is uncertain to happen'. Formally, we define:

Definition 3.3. A **contradiction interaction** between recommendations $R, R' \in \mathbb{R}$ is a tuple (R, R', μ), where $\mu \in \{\Box, \Diamond\}$ is the *modal strength* of the interaction.

$\mathbb{I}$ denotes the set of all contradiction interactions given $\mathbb{R}$.

Example 3.4. The recommendations R_1 and R_2 from Example 3.2 are in a contradiction interaction, as they recommend opposite actions.[2] We assume that $\mathbb{I} = \{(R_1, R_2, \Box)\}$.

The possibility to identify interactions gives rise to the following notion of *contradiction-free* sets of recommendations.

Definition 3.5. A set $\mathbb{R}' \subseteq \mathbb{R}$ is **contradiction-free** iff there is no contradiction interaction $(R_i, R_j, \mu) \in \mathbb{I}$ with $R_i, R_j \in \mathbb{R}'$.

Intuitively, contradiction-free sets of recommendations consist of recommendations that are safe to follow without the risk of performing incompatible actions.

[2]Note well that a hierarchy of actions is assumed in [42, p. 79] to obtain interactions. For instance, the action to administer NSAID subsumes both actions to administer Aspirin and Ibuprofen. This hierarchy is not important for our purposes.

Example 3.6. The set $\mathbb{R} = \{R_1, R_2\}$ from Example 3.2 is not contradiction-free, for $(R_1, R_2, \square) \in \mathbb{I}$, as in Example 3.4. Clearly, $\{R_1\}$ and $\{R_2\}$ themselves are contradiction-free.

Our representation of recommendations and interactions as afforded by the TMR model will contribute to our approach meeting the 1^{st} and the 3^{rd} Ariadne principles as laid down in Section 2.

3.2 Context

Recommendations $\mathbb{R}$ and interactions $\mathbb{I}$ amount only to representation of guidelines, but not reasoning with them. In particular, they give a patient-agnostic representation, while the reasoning happens with patient-specific information.

Example 3.7. Consider $\mathbb{R} = \{R_1, R_2\}$ and $\mathbb{I} = \{(R_1, R_2, \square)\}$ as in Examples 3.2 and 3.4. Intuitively, for a generic patient, NSAID – e.g. Aspirin – should be administered. If, however, the patient exhibits *Gastro. Bleeding*, then R_1 and R_2 are in conflict and there are arguments for both administering and not administering Aspirin.

The patient information can be understood as the *context* in which reasoning happens (see e.g. [32]). To resolve the conflict in Example 3.7, one could administer a different NSAID, such as Ibuprofen. However, in more complicated situations such alternatives may not be available. In those situations, *preferences* may be a part of the context that help to resolve the conflicts argumentatively.

Example 3.8. Continuing Example 3.7, suppose that only Aspirin is available. The patient may insist that medication should be given to them, thus preferring taking Aspirin over not taking it, whence only R_1 should be followed. On the other hand, if the patient expresses no preferences, the clinician's priorities may come into play. For instance, the clinician may deem not increasing the risk of gastrointestinal bleeding more important than decreasing blood coagulation, whence only R_2 would be followed.

Thus, the context includes not only the patient's state, but also various preferences. For instance: a) the patient may prefer one course of action over another; b) the clinician may prioritise treatments in accordance with patient-centric goals and their importance. The TMR model however does not afford representation of such preferences, just as it does not afford representation of patient-specific conditions. One of our tasks is to augment the representation of recommendations and interactions with the context of a patient so as to enable patient-centric reasoning with clinical guidelines. For this purpose, we define the context pertaining to patient information with respect to recommendations as follows.

Definition 3.9. The **context** (of a fixed but otherwise arbitrary patient) is a tuple $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$ with: the patient's state $\mathcal{S}$, the patient-centric goals $\mathcal{G}$, the (patient's) preferences $\leqslant$ over actions, the (clinician's) priorities $\preccurlyeq$ over goals.[3]

In the rest of the paper we assume that a context is compatible with given recommendations in the following sense: the patient's state $\mathcal{S}$ matches some of the properties within recommendations; the goals $\mathcal{G}$ match the (un)desired effects on those properties; the patient's preferences are (represented by a preorder) over the recommended actions or recommendations; the clinician's priorities are (represented by a total preorder) over the effects on the patient's state. We make this precise in Section 4.3.

Example 3.10. Building on Examples and 3.7 and 3.8, the context of the patient can be given by $\mathcal{S} = \{Gastro.\ Bleeding\}$, $\mathcal{G} = \{decrease\ Blood\ Coag., \text{not}\ increase\ Gastro.\ Bleeding\}$,[4] $R_2 < R_1$,[5] and *decrease Blood Coag.* $\prec$ not *increase Gastro. Bleeding*.

The elements together form a context for the application of recommendations and ground them to a particular setting. The context of a patient will contribute to our approach meeting the 2^{st} and the 3^{rd} Ariadne principles put forward in Section 2.

4 REASONING WITH GUIDELINES

We will use guideline recommendations, their interactions and contexts to construct argumentation frameworks for an agent to reason and resolve conflicts among recommendations, given patient-specific conditions, patient-centric goals and various preferences. Specifically, we will use ABA$^+$ frameworks for assumption-based reasoning with guidelines and patient's preferences over recommendations. We will then augment ABA$^+$ to ABA$^+$G for goal-driven reasoning with guidelines and clinician's priorities over goals.

4.1 ABA$^+$ Background

We provide the background for ABA$^+$ following [7, 8].

An **ABA$^+$ framework** is a tuple $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$, where:

- $(\mathcal{L}, \mathcal{R})$ is a deductive system with $\mathcal{L}$ a language and $\mathcal{R}$ a set of rules of the form $\varphi_0 \leftarrow \varphi_1, \dots, \varphi_m$ with $m \geqslant 1$, or of the form $\varphi_0 \leftarrow \top$, where $\varphi_i \in \mathcal{L}$ for $i \in \{0, \dots, m\}$ and $\top \notin \mathcal{L}$; φ_0 is the *head* or *conclusion*, and $\varphi_1, \dots, \varphi_m$ the *body* of the rule; $\varphi_0 \leftarrow \top$ is said to have an empty body and called a *fact*;
- $\mathcal{A} \subseteq \mathcal{L}$ is a non-empty set of *assumptions*;
- $^- : \mathcal{A} \to \mathcal{L}$ is a total map: for $\alpha \in \mathcal{A}$, $\overline{\alpha}$ is referred to as the *contrary* of α;
- $\leqslant$ is a preorder (i.e. reflexive and transitive order) on $\mathcal{A}$, called a *preference relation*.

For $\alpha, \beta \in \mathcal{A}$, $\alpha \leqslant \beta$ means that β is at least as preferred as α, and $\alpha < \beta$ means that α is strictly less preferred than β.

Throughout, we assume a fixed but otherwise arbitrary ABA$^+$ framework $\mathcal{F} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$, unless else specified.

Assumptions in ABA$^+$ represent arguable information. For instance, assumptions can represent the applicability of, or an agent's willingness to follow, a recommendation. In such a case, preferences in ABA$^+$ can represent the willingness to follow recommendations.

We next give notions of arguments and attacks in ABA$^+$.

An **argument for conclusion** $\varphi \in \mathcal{L}$ **supported by** $A \subseteq \mathcal{A}$ **and** $R \subseteq \mathcal{R}$, denoted $A \vdash^R \varphi$, is a finite tree with: the root labelled by φ; leaves labelled by $\top$ or assumptions, with A being the set of all such assumptions; the children of non-leaves ψ labelled by the elements of the body of some ψ-headed rule in $\mathcal{R}$, with R being the set of all such rules. $A \vdash \varphi$ abbreviates $A \vdash^R \varphi$ with some $R \subseteq \mathcal{R}$.

For $A, B \subseteq \mathcal{A}$, A $<$**-attacks** B, denoted $A \rightsquigarrow_< B$, iff:

[3] Following Ariadne principles, we distinguish between preferences over actions and priorities over goals for ease of reference.

[4] not is purely syntactic, representing the desire to avoid the effect on the property brought about by the action.

[5] As usual, the strict (asymmetric) counterpart $<$ of a preorder $\leqslant$ is given by $\alpha < \beta$ iff $\alpha \leqslant \beta$ and $\beta \nleqslant \alpha$, for any α and β. We assume this for all preorders in this paper.

a) either there is an argument $A' \vdash \overline{\beta}$, for some $\beta \in B$, supported by $A' \subseteq A$, and $\nexists \alpha' \in A'$ with $\alpha' < \beta$;
b) or there is an argument $B' \vdash \overline{\alpha}$, for some $\alpha \in A$, supported by $B' \subseteq B$, and $\exists \beta' \in B'$ with $\beta' < \alpha$.

The intuition here is that A $<$-attacks B if a) either A argues contra something in B by means of no inferior elements (*normal attack*), b) or B argues contra something in A but with at least one inferior element (*reverse attack*).

If A does not $<$-attack B, we may write $A \not\rightsquigarrow_< B$. Note that without preferences, an attack from one set of assumptions to another boils down to the former set deducing the contrary of some assumption in the latter set.

We next give notions used to define ABA$^+$ semantics.

Let $A \subseteq \mathcal{A}$. The *conclusions of* A is the set of sentences $Cn(A) = \{\varphi \in \mathcal{L} \ : \ \exists\, A' \vdash \varphi,\ A' \subseteq A\}$ concluded by (arguments supported by subsets of) A. We say A is *closed* if $A = Cn(A) \cap \mathcal{A}$, i.e. A contains all assumptions it concludes. We say $\mathcal{F}$ is *flat* if every $A \subseteq \mathcal{A}$ is closed. We assume ABA$^+$ frameworks to be flat in this paper.

Further: A is *$<$-conflict-free* if $A \not\rightsquigarrow_< A$; also, A *$<$-defends* $A' \subseteq \mathcal{A}$ if $\forall\, B \subseteq \mathcal{A}$ with $B \rightsquigarrow_< A'$ we have $A \rightsquigarrow_< B$; and A is *$<$-admissible* if it is $<$-conflict-free and $<$-defends itself. We consider one ABA$^+$ semantics: a set $E \subseteq \mathcal{A}$ of assumptions is a **$<$-preferred extension** of $\mathcal{F} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant)$ if E is $\subseteq$-maximally $<$-admissible.

4.2 ABA$^+$G: ABA$^+$ with Goals

We extend ABA$^+$ with a mechanism to distinguish among preferred extensions based on goals fulfilled. Oliveira et al. introduce goal seeking mechanisms in structured argumentation to rank argument extensions according to their relative priorities [26]. We import this goal-driven reasoning into ABA$^+$ to define ABA$^+$G, and thus cover the important aspect of reasoning with patient-centric goals.

Definition 4.1. An **ABA$^+$G argumentation framework** is a tuple $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant, \mathcal{G}, \preccurlyeq)$, where $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant)$ is an ABA$^+$ framework and

- $\mathcal{G} \subseteq \mathcal{L}$ is a finite set of **goals** such that $\forall\ \theta \in \mathcal{G}$, there exists $\theta \leftarrow \varphi_1, \ldots, \varphi_m$ with $m \leqslant 1$ in $\mathcal{R}$;
- $\preccurlyeq$ is a total preorder on $\mathcal{G}$, denoting **priorities** over goals; for $\theta, \chi \in \mathcal{G}$, $\theta \preccurlyeq \chi$ means χ is as important as θ.

In what follows, $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant, \mathcal{G}, \preccurlyeq)$ is a fixed but otherwise arbitrary ABA$^+$G framework, unless said otherwise.

In ABA$^+$G concluding goals amounts to fulfilling them. We hence define (preferred) *goal extensions* in terms of goal-conclusions thus:

Definition 4.2. Let E be a $<$-preferred extension of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant)$. Then $\mathcal{G}_E = Cn(E) \cap \mathcal{G}$ is a **goal extension** of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant, \mathcal{G}, \preccurlyeq)$.

In other words, a goal extension consists of the goals concluded by a $<$-preferred extension. We use priorities over goals to rank goal extensions and define ABA$^+$G semantics:

Definition 4.3. Let $\mathbb{G}$ be the set of goal extensions. The *goal extension ordering* $\trianglelefteq_{\mathbb{G}}$ over $\mathbb{G}$ is given by

$$\mathcal{G}_A \trianglelefteq_{\mathbb{G}} \mathcal{G}_B \text{ iff } \exists \theta \in \mathcal{G}_B \setminus \mathcal{G}_A \text{ with } \chi \preccurlyeq \theta\, \forall \chi \in \mathcal{G}_A \setminus \mathcal{G}_B.$$

$\mathcal{G} \in \mathbb{G}$ is a **top goal extension** iff $\nexists \mathcal{G}' \in \mathbb{G}$ such that $\mathcal{G} \vartriangleleft_{\mathbb{G}} \mathcal{G}'$.

Note that $\trianglelefteq_{\mathbb{G}}$ is a total preorder, as $\preccurlyeq$ is a total preorder. Intuitively, $\mathcal{G}_A \trianglelefteq_{\mathbb{G}} \mathcal{G}_B$ means that $\mathcal{G}_B$ is at least as 'good' as $\mathcal{G}_A$. The underlying principle behind ordering goal extensions is trying to fulfill goals according to their importance.

A top goal extension admits no strictly 'better' goal extension. Intuitively, a $<$-preferred ABA$^+$ extension inducing a top goal extension yields the best reasoning outcome.

This choice of ordering is motivated by the requirements of a patient management setting, within which priorities over goals may convey a sense of urgency and severity that must be addressed when reasoning. Hence, we assume that an agent should always aim to fulfill the top preferred goals, regardless of the goals with lower priorities. In general, preference aggregation is a rich and complex area of research. Other orderings could be applied, see e.g. [16] for a comparison of various orderings, but we chose the above one in accordance to our interpretation of priorities over goals.

4.3 Reasoning in ABA$^+$G

We now introduce the representation in ABA$^+$G of TMR instances, interactions and context.

4.3.1 Intuition. At a high-level: assumptions will represent (the defeasible applicability of) recommendations, the corresponding actions and their effects on properties will be modelled via rules, and the deontic strength will determine both whether the actions and their consequences are sought after or not. The context will be modelled via facts representing patient's state, goals matching the effects of actions, patient's preferences over assumptions and clinician's priorities over goals.

For a step by step illustration, we use recommendations $\mathbb{R} = \{R_1, R_2\}$ and interactions $\mathbb{I} = \{(R_1, R_2, \Box)\}$ as in Example 3.7. First, $R_1, R_2 \in \mathcal{A}$ represent the (defeasible applicability of) recommendations. The following rules then represent the actions recommended (or not) by R_1 and R_2:

1. *Adm. NSAID* $\leftarrow R_1$;
2. not *Adm. Aspirin* $\leftarrow R_2$.

The following rules model the effects the actions *Adm. NSAID* and *Adm. Aspirin* bring about:

3. *decrease Blood Coag.* $\leftarrow$ *Adm. NSAID*;
4. *increase Gastro. Bleeding* $\leftarrow$ *Adm. Aspirin.*

As R_2 recommends not *Adm. Aspirin*, the following rule represents the effect to be avoided by following R_2:

5. not *increase Gastro. Bleeding* $\leftarrow$ not *Adm. Aspirin.*

Now, R_1 and R_2 are in contradiction with *Adm. NSAID* and *Adm. Aspirin* recommended positively and negatively, respectively. Thus, R_2 can be argued against on the basis of R_1 and the presence of the contradiction. However, R_1 can be similarly argued against on the basis of R_2 and the presence of the contradiction, but *only as long as* the given patient has *Gastro. Bleeding*. Therefore, we have:

6. $\overline{R_2} \leftarrow R_1, int_{1,2}$;
7. $\overline{R_1} \leftarrow R_2, int_{1,2}$, *Gastro. Bleeding.*

Here, $\overline{R_1}$ and $\overline{R_2}$ are the contraries of R_1 and R_2, respectively, and $int_{1,2} \in \mathcal{L}$ represents $(R_1, R_2, \mu) \in \mathbb{I}$. These rules say that:

R_2 should not be followed if
(i) R_1 is followed, and
(ii) R_1 and R_2 are in contradiction;

R_1 should not be followed if
(i) R_2 is followed,
(ii) R_1 and R_2 are in contradiction, *and also*

(iii) the condition *Gastro. Bleeding* is present.

This is in accordance with the desirable reading of interactions as in Section 3 and in [42, p. 91].

The interaction's modal strength μ determines whether it is an assumption (i.e. could be argued about) or a fact (i.e. sure to happen):
a) if $\mu = \Box$, let $int \leftarrow \top \in \mathcal{R}$;
b) if $\mu = \Diamond$, let $int \in \mathcal{A}$.
In our example, $\mu = \Box$, so we have:

8. $int_{1,2} \leftarrow \top$.

Given context $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$, the patient's state $\mathcal{S}$ yields properties/initial value-property pairs as facts. With context from Example 3.10, *Gastro. Bleeding* $\in \mathcal{S}$ yields:

9. *Gastro. Bleeding* $\leftarrow \top$.

Lastly, as in Example 3.10, goals $\mathcal{G}$ represent (un-)desired effects on properties, patient's preferences $\leqslant$ are over recommendations as assumptions and clinician's priorities are over goals.

4.3.2 Formalisation. Formally, mapping recommendations, interactions and context to ABA$^+$G goes as follows.

Definition 4.4. Given recommendations $\mathbb{R}$, interactions $\mathbb{I}$ and context $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$, **the ABA$^+$G patient framework** is defined as $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$, where:

- $\mathcal{A} = \{R : (R, A, \delta, \mathcal{P}, \mathcal{E}, \mathcal{V}, \mathcal{C}) \in \mathbb{R}\} \cup \{int_{i,j} : (R_i, R_j, \Diamond) \in \mathbb{I}\}$ consists of assumptions representing recommendations and uncertain interactions;
- $\mathcal{R}_a = \mathcal{R}_a^+ \cup \mathcal{R}_a^-$ consists of rules representing actions associated to recommendations, where
  - $\mathcal{R}_a^+ = \{A \leftarrow R : R \in \mathbb{R}, \delta \geqslant 0\}$,
  - $\mathcal{R}_a^- = \{\text{not } A \leftarrow R : R \in \mathbb{R}, \delta < 0\}$;[6]
- $\mathcal{R}_e = \mathcal{R}_e^+ \cup \mathcal{R}_e^-$ consists of rules representing effects brought about by actions, where
  - $\mathcal{R}_e^+ = \{E^k P^k \leftarrow A : R \in \mathbb{R}, E^k \in \mathcal{E}, P^k \in \mathcal{P}, \delta \geqslant 0\}$,
  - $\mathcal{R}_e^- = \{\text{not } E^k P^k \leftarrow \text{not } A : R \in \mathbb{R}, E^k \in \mathcal{E}, P^k \in \mathcal{P}, \delta < 0\}$;
- $\mathcal{R}_s = \{vP \leftarrow \top : vP \in \mathcal{S}\}$ consists of facts representing the patient's state $\mathcal{S}$ in terms of properties and their values, where $\mathcal{S} \subseteq \bigcup_{R \in \mathbb{R}} \{v^k P^k : P^k \in \mathcal{P}, v^k \in \mathcal{V}\}$;
- $\mathcal{R}_c = \mathcal{R}_c^+ \cup \mathcal{R}_c^-$ consists of rules representing (contradiction) interactions between recommendations, where
  - $\mathcal{R}_c^+ = \{\overline{R_j} \leftarrow R_i, int_{i,j} : (R_i, R_j, \mu) \in \mathbb{I}, \delta_i \geqslant 0\}$,
  - rules in $\mathcal{R}_c^- = \{\overline{R_i} \leftarrow R_j, int_{i,j}, v_j^k P_j^k : (R_i, R_j, \mu) \in \mathbb{I}, \delta_j < 0, P_j^k \in \mathcal{P}_j, v_j^k \in \mathcal{V}_j, c_j^k \in \mathcal{C}_j, c_j^k = -\}$ take into account presence of negatively affected conditions;
- $\mathcal{R} = \mathcal{R}_a \cup \mathcal{R}_e \cup \mathcal{R}_s \cup \mathcal{R}_c \cup \{int_{i,j} \leftarrow \top : (R_i, R_j, \Box) \in \mathbb{I}\}$ consists of rules defined above and rules representing interactions that are sure to happen;
- $\leqslant$ is a preorder over $\mathcal{A}$;
- $\mathcal{G} = \mathcal{G}^+ \cup \mathcal{G}^-$ satisfies
  - $\mathcal{G}^+ \subseteq \bigcup_{R \in \mathbb{R}} \{E^k P^k : P^k \in \mathcal{P}, E^k \in \mathcal{E}\}$,
  - $\mathcal{G}^- \subseteq \bigcup_{R \in \mathbb{R}} \{\text{not } E^k P^k : P^k \in \mathcal{P}, E^k \in \mathcal{E}\}$,
- $\preccurlyeq$ is a total preorder over $\mathcal{G}$;
- By convention, $\mathcal{L}$ and $^-$ are implicit from $\mathcal{A}$ and $\mathcal{R}$ as follows: unless $\overline{x}$ appears in either $\mathcal{A}$ or $\mathcal{R}$, it is different from the sentences appearing in $\mathcal{A}$ or $\mathcal{R}$; thus, $\mathcal{L}$ consists of all the sentences appearing in $\mathcal{R}$, $\mathcal{A}$ and $\{\overline{\alpha} : \alpha \in \mathcal{A}\}$.

[6]not is purely syntactic (see footnote 4).

Regarding interactions and rules in $\mathcal{R}_c$, suppose recommendations R_i and R_j are in contradiction with actions A_i and A_j recommended positively ($\delta_i > 0$) and negatively ($\delta_j < 0$), respectively. On the one hand, R_j can be argued against on the basis of R_i and the presence of the interaction. On the other hand, R_i can be similarly argued against on the basis of R_j and the presence of the interaction, but *only as long as* a given patient will have some condition affected by A_j that contributes negatively to the patient's well-being. Thus, we take into account any property $P \in \mathcal{P}_j$ with initial value $v \in \mathcal{V}_j$ and contribution $- = c \ni \mathcal{C}_j$. When the initial value v of P is indeterminate ?, we use only P.

4.3.3 Properties. Modelling recommendations and interactions argumentatively allows to exploit properties of ABA$^+$ to ensure desirable features of our approach. Specifically, the $<$-preferred extensions in ABA$^+$G patient frameworks are contradiction-free (Definition 3.5) as sets of recommendations (recall that we identify a recommendation with its name, see remark after Definition 3.1):

Theorem 4.5 (**Interaction Theorem**). *For a $<$-preferred extension E of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$ in $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$, $E \cap \mathbb{R}$ is a contradiction-free set of recommendations.*

Proof. Suppose $E \cap \mathbb{R}$ is not contradiction-free. Then there is $(R_i, R_j, \mu) \in \mathbb{I}$ with $R_i, R_j \in E$. But as $\overline{R_j} \leftarrow R_i, int_{i,j} \in \mathcal{R}$ and $int_{i,j}$ is either a fact or an $<$-unattacked assumption, we find $E \rightsquigarrow_< E$. This contradicts $<$-conflict-freeness of E. □

Thus, top goal extensions (induced by $<$-preferred extensions) in ABA$^+$G are guaranteed to yield goals achievable without the risk of performing incompatible actions.

Another property states that if the patient expresses preferences over all recommendations, then the most preferred non-conflicting recommendations will be followed:

Theorem 4.6 (**Preferences Theorem**). *Let $\leqslant$ be total over $\mathcal{A} \cap \mathbb{R}$ and the set $\mathbb{R}' = \{R \in \mathcal{A} : \nexists R' \in \mathcal{A} \text{ with } R < R'\}$ of the most preferred recommendations be contradiction-free. Then $\mathbb{R}' \subseteq E$ for every $<$-preferred extension E of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$ in $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$.*

Proof. Any $R \in \mathbb{R}'$ is $\leqslant$-maximal, so $\{R\}$ is $<$-unattacked. As all $<$-attacks come from by singleton sets, every $<$-preferred extension contains all the $<$-unattacked sets of assumptions, including $\mathbb{R}'$. □

Theorems 4.5 and 4.6 ensure that ABA$^+$G meets the three Ariadne principles of *interaction assessment*, *prioritisation and patient preferences* and *individualised management* when applied to patient-centric reasoning with conflicting medical recommendations.

4.3.4 Illustration. We exemplify our formalisation with a case study from [42], focusing on contradiction interactions between breast cancer (BC) and hypertension (HT) guidelines, and using the pertinent parts of the information (given in full in [42, p. 87, Figure 5, p. 90, Table 9, p. 91, Table 10]). Our results will be in agreement with the informal discussion on the case study in [42].

Example 4.7. We assume a merged BC and HT guideline with:

- (R_8, *High Int. Exercise*, *should not*, ⟨*Blood Press.*⟩, ⟨*increase*⟩, ⟨?⟩, ⟨−⟩),
- (R_4, *Exercise*, *must not*, ⟨*Body Temp.*⟩, ⟨*increase*⟩, ⟨*high*⟩, ⟨−⟩),
- (R_3, *Low Int. Exercise*, *should*, $\mathcal{P}_3, \mathcal{E}_3, \mathcal{V}_3, \mathcal{C}_3$),

- $(R_2, \textit{Std. Exercise}, \textit{should}, \mathcal{P}_3 \cup \langle \textit{Lymphedema risk} \rangle, \mathcal{E}_3 \cup \langle \textit{increase} \rangle, \mathcal{V}_3 \cup \langle \textit{present} \rangle, \mathcal{C}_3 \cup \langle - \rangle)$, where
  - $\mathcal{P}_3 = \langle \textit{Fatigue}, \textit{Fitness}, \textit{Pain} \rangle$,
  - $\mathcal{E}_3 = \langle \textit{decrease}, \textit{decrease}, \textit{decrease} \rangle$,
  - $\mathcal{V}_3 = \langle \textit{high}, \textit{high}, \textit{high} \rangle$,
  - $\mathcal{C}_3 = \langle +, +, + \rangle$.

For instance, R_8 says that one *should not* do *High Int. Exercise*, because it negatively contributes by increasing *Blood Press.*; R_3 says that one *should* do *Low Int. Exercise*, because it positively contributes to decreasing *Fatigue*, *Fitness* and *Pain* from *high* values.

Thus, $\mathbb{R} = \{R_2, R_3, R_4, R_8\}$ and the interactions identified are $\mathbb{I} = \{(R_2, R_4, \square),\ (R_3, R_4, \square),\ (R_2, R_8, \square)\}$.

To illustrate reasoning with patient-specific conditions, goals and preferences, we assume, in addition to the case study of [42], *patient A*. Let patient A exhibit increased *Blood Press.* (indeterminate value), and in addition have *high Body Temp.* Suppose patient A has also expressed an overall preference for not doing high intensity exercise: $R_2 < R_8$, $R_3 < R_8$, $R_4 < R_8$.

After discussing with patient A, the clinician elaborates a list of patient-centric goals, thus:

- $\mathcal{G} = \{$*decrease Pain*, not *increase Blood Press.*, *decrease Fatigue*, not *increase Body Temp.*$\}$.

Note that not all properties (and effects) from recommendations need be included in $\mathcal{G}$: for instance, *Lymphedema risk* is not concerned. The prioritisation of goals may be motivated by several criteria. Their specification is outside the scope of this work, but an example is the severity of a condition over the property it is associated with. For patient A, the pain level from BC is a significant considerable obstacle to daily life, impeding normal routine. Additionally, the clinician is concerned with the patient's high blood pressure. Thus, *decrease Pain* is the strictly most important goal, followed by not *increase Blood Press.*, which is followed by the equally important *decrease Fatigue* and not *increase Body Temp.* Then $\preccurlyeq$ over $\mathcal{G}$ is defined as follows: *decrease Fatigue* $\preccurlyeq$ not *increase Body Temp.*; not *increase Body Temp.* $\preccurlyeq$ *decrease Fatigue*; *decrease Fatigue* $\prec$ not *increase Blood Press.* $\prec$ *decrease Pain*; and visualised thus:

decrease Pain
|
not *increase Blood Press.*
/ \
decrease Fatigue not *increase Body Temp.*

The associated ABA$^+$G framework $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^{-}, \leqslant, \mathcal{G}, \preccurlyeq)$:

- $\mathcal{A} = \{R_2, R_3, R_4, R_8\}$,
- $\mathcal{R} = \{$*Std. Exercise* $\leftarrow R_2$, *Low Int. Exercise* $\leftarrow R_3$,
  not *Exercise* $\leftarrow R_4$, not *High Int. Exercise* $\leftarrow R_8\} \cup \{$
  increase Body Temp. $\leftarrow$ *Std. Exercise*,
  decrease Fatigue $\leftarrow$ *Std. Exercise*,
  decrease Pain $\leftarrow$ *Std. Exercise*,
  decrease Fatigue $\leftarrow$ *Low Int. Exercise*,
  decrease Pain $\leftarrow$ *Low Int. Exercise*,
  not *increase Blood Press.* $\leftarrow$ not *High Int. Exercise*,
  not *increase Body Temp.* $\leftarrow$ not *Exercise*$\} \cup$
  $\{$*Blood Press.* $\leftarrow \top$, *high Body Temp.* $\leftarrow \top\} \cup$
  $\{\overline{R_4} \leftarrow R_2, int_{2,4},\ \overline{R_2} \leftarrow R_4, int_{2,4}, \textit{high Body Temp.},$
  $\overline{R_4} \leftarrow R_3, int_{3,4},\ \overline{R_3} \leftarrow R_4, int_{3,4}, \textit{high Body Temp.},$
  $\overline{R_8} \leftarrow R_2, int_{2,8},\ \overline{R_2} \leftarrow R_8, int_{2,8}, \textit{Blood Press.}\} \cup$
  $\{int_{2,4} \leftarrow \top,\ int_{3,4} \leftarrow \top,\ int_{2,8} \leftarrow \top\}$,
- $\leqslant$, $\mathcal{G}$ and $\preccurlyeq$ as above (and $\mathcal{L}$ and ${}^{-}$ as per convention).

All contradiction interactions are triggered, giving arguments: $\{R_2\} \vdash \overline{R_4}$, $\{R_2\} \vdash \overline{R_8}$, $\{R_4\} \vdash \overline{R_2}$, $\{R_4\} \vdash \overline{R_3}$, $\{R_3\} \vdash \overline{R_4}$, $\{R_8\} \vdash \overline{R_2}$. These indicate which recommendations are contradicting which other ones. Then, patient preferences help to determine the 'stronger' arguments and (non-) $<$-attacks: $\{R_2\} \rightsquigarrow_< \{R_4\}$; $\{R_4\} \rightsquigarrow_< \{R_2\}$; $\{R_3\} \rightsquigarrow_< \{R_4\}$; $\{R_4\} \rightsquigarrow_< \{R_3\}$; $\{R_8\} \rightsquigarrow_< \{R_2\}$; but $\{R_2\} \not\rightsquigarrow_< \{R_8\}$. We thus see that, in particular, R_2 suggesting *Std. Exercise*, contradicting R_8 but being less preferred, does not stand as an argument against *High Int. Exercise* suggested by R_8.

ABA$^+$ semantics then resolves the conflicts. Briefly, as R_3 and R_4 are mutually contradicting, but each non-interacting with R_8, they can be followed alongside R_8, which itself kicks out R_2. Thus, $\{R_3, R_8\}$, $\{R_4, R_8\}$ are $<$-preferred extensions. The former suggests *Low Int. Exercise* and advises against *High Int. Exercise*; the latter urges not to exercise at all. The corresponding goal extensions:

- $\mathcal{G}_{\{R_3,R_8\}} = \{$*decrease Fatigue*, not *increase Blood Press.*, *decrease Pain*$\}$;
- $\mathcal{G}_{\{R_4,R_8\}} = \{$*decrease Fatigue*, not *increase Body Temp.*, *decrease Pain*$\}$.

In both goal extensions $\mathcal{G}_{\{R_3,R_8\}}$ and $\mathcal{G}_{\{R_4,R_8\}}$ the most important goal *decrease Pain* is fulfilled, and so is *decrease Fatigue*. But only $\mathcal{G}_{\{R_3,R_8\}}$ fulfills the second most important goal not *increase Blood Press.*, so it is strictly better: $\mathcal{G}_{\{R_4,R_8\}} \lhd_{\mathbb{G}} \mathcal{G}_{\{R_3,R_8\}}$. Consequently, $\mathcal{G}_{\{R_3,R_8\}}$ is the top goal extension. Accordingly, R_3 (*Low Int. Exercise*) and R_8 (not *High Int. Exercise*) should be followed.

We showed how ABA$^+$G allows for assumption-based, goal-driven and preference-respecting reasoning with TMR recommendations and interactions, taking into account patient-specific conditions, preferences over recommendations and priorities over goals.

5 RELATED WORK

Argumentation (with or without preferences) has been successfully applied in health care (see e.g. [3, 21] for overviews). For instance, in [15], evidence from clinical trials is manually extracted from guidelines and synthesised to form arguments for treatment superiority, with attacks among arguments with conflicting claims. Based on treatment outcome indicators and the importance of evidence, user-specified preferences over arguments and argumentation semantics are used to identify the acceptable arguments. The focus is determining superiority among treatments, not concerning guideline recommendations or conflict resolution among those. Instead, argument aggregation for reasoning with guidelines is used in [14]. There, arguments correspond to statements in guidelines and, for a single specified goal, confidence of arguments is aggregated to identify the acceptable arguments. The focus is enacting recommendations from a single guideline, rather than reasoning with conflicting recommendations from multiple guidelines. Other works, e.g. [11, 29, 35], integrate argumentation with preferences to help clinicians to construct, exchange and evaluate arguments for and against decisions, rather than to reason with guidelines.

The recent CONSULT project [18] applies argumentation to reason with guidelines and patient preferences for managing post-stroke patients. Kokciyan et al. manually represent guidelines in first-order logic (FOL) and use argument schemes [40] to construct arguments. We believe that using FOL for guideline representation is ad-hoc, and instead use the well-established TMR model to represent guideline recommendations and identify their interactions, which we then map into ABA$^+$G. Further, Kokciyan et al. use argumentation with preferences modelled as attacks on attacks [22] to resolve conflicts among recommendations. We instead incorporate preferences directly in the construction of attacks in ABA$^+$G. Importantly, this aspect and the ability to model and reason with goals allows us to meet Ariadne principles of patient management. We leave formal comparison with [18] for future work.

[41] is a recent non-argumentative approach to reasoning with interacting guidelines, patient conditions and preferences. There, guideline recommendations are represented as actionable graphs and mapped into first-order logic (FOL) rules, while patient conditions and preferences are represented as FOL revision operators. Reasoning (guideline mitigation) amounts to applying revision operators to account for patient-specific conditions and preferences, and then finding models of the resulting FOL theory. Our approach is different in both knowledge representation—TMR model is richer than the mitigation-specific FOL, and computation mechanism—model finding is undecidable as opposed to finding preferred extensions. We also believe argumentation-based reasoning to be more transparent, as one can inspect the arguments, attacks among them and their interplay with preferences, in contrast to interpreting workings and results of a FOL theorem prover.

Other approaches to reasoning with guidelines (see [28, 31] for overviews) focus on execution of single guidelines, e.g. [20, 33], or identification of incompatibilities among guidelines, e.g. answer set programming is used in [34] to check temporal conformance; statistical preference learning is used in [36] to identify inconsistencies in antibiotic therapy guidelines. Yet other works concern preference elicitation to facilitate shared (clinician-patient) decision making. In particular, Sacchi et al. incorporate patients' preferences in terms of QALYs, utilities and costs into the shared decision making model [32]. In effect, they propose a framework that supports patient preference elicitation and integrates them with patient health record to feed into decision models (particularly, decision trees) so as to facilitate shared (clinician-patient) decision making. This allows to better inform both the clinician and the patient about the alternatives, but does not afford automatic resolution of interacting (e.g. conflicting) recommendations. It would be interesting to see how this could inform knowledge representation in our approach.

Goal-driven argumentative decision making (possibly with preferences) has been explored, e.g. [1, 10, 24, 43]. The settings there do not apply to reasoning with guidelines. As for goal-driven argumentative decision-making, the approach of [1] concerns general multiple criteria decision making in argumentation with preferences via reasoning backwards from goals to arguments. A follow-up application-specific approach of [24] affords goal-driven argumentative documentation, analysis and making of decisions. ABA$^+$G differs from these approaches particularly in the direction of reasoning—from arguments to goals, which is more similar to assumption-based reasoning with goals and preferences as in [10]; as well as in using preferences (over goals) to select among extensions, as in e.g. [2, 39]. It would be interesting to investigate the formal relationships with all these works in the future.

We note that an argumentative approach with context was recently proposed in [43], where context rules and primitives involving patient state properties are used to assert defeasibility of logical implications between decisions, attributes, and goals. Thus, context-sensitivity is an important and desirable property in both medical and argumentative settings, and we addressed it in this work.

6 CONCLUSIONS AND FUTURE WORK

We proposed ABA$^+$ to reason with guidelines and patient context. We mapped guideline representation as TMR model to ABA$^+$, incorporated in ABA$^+$ patient-specific conditions and preferences, and augmented ABA$^+$ to ABA$^+$G so as to account for patient-centric goals and their importance. ABA$^+$G yields contradiction-free recommendations and associated achievable goals while respecting the context of the patient. Our approach meets Ariadne principles: *interaction assessment* is ensured by Theorem 4.5 (contradiction-free recommendations); *prioritisation and patient preferences* is ensured by Definition 4.3 (extensions fulfill the most important goals) and Theorem 4.6 (most preferred non-conflicting recommendations are followed); *individualised management* is ensured collectively by the above and the use of the patient context in ABA$^+$G (Definition 4.4). To the best of our knowledge, our work is unique in establishing a relationship between features of argumentative reasoning and principles of patient management. This hints at the adequacy of structured argumentation for this type of task, which we believe is important given the difficulty (both in terms of time and resources) of large-scale practical evaluations in a real setting.

In addition to several future work directions mentioned in Section 5, we will extend our work to other interaction types identified in [42]: repetition, alternative, etc. We will also aim to incorporate various numerical measures from TMR, such as belief strength. This may yield additional preferences, and we will study how multiple types of possibly conflicting preferences can be simultaneously integrated in ABA$^+$G. Preference elicitation is a vast problem by itself, and we will explore integration with the relevant works, e.g. [32]. Last but not least, argumentation is well-suited for explanations, see e.g. [3, 23], and we will study both the well-established and novel ABA$^+$ mechanisms to explain to the patient or the clinician how and why ABA$^+$G arrives at the final recommendations.

Acknowledgements. Kristijonas Čyras was supported by the EPSRC project **EP/P029558/1** ROAD2H: *Resource Optimisation, Argumentation, Decision Support and Knowledge Transfer to Create Value via Learning Health Systems.* Tiago Oliveira was supported by JSPS KAKENHI Grant Number **JP18K18115**.

Data access statement: No new data was collected in the course of this research.

REFERENCES

[1] Leila Amgoud and Henri Prade. 2009. Using arguments for making and explaining decisions. *Artificial Intelligence* 173, 3-4 (2009), 413–436. https://doi.org/10.1016/j.artint.2008.11.006

[2] Leila Amgoud and Srdjan Vesic. 2014. Rich Preference-Based Argumentation Frameworks. *International Journal of Approximate Reasoning* 55, 2 (2014), 585–606. https://doi.org/10.1016/j.ijar.2013.10.010

[3] Katie Atkinson, Pietro Baroni, Massimiliano Giacomin, Anthony Hunter, Henry Prakken, Chris Reed, Guillermo Ricardo Simari, Matthias Thimm, and Serena Villata. 2017. Towards Artificial Argumentation. *AI Magazine* 38, 3 (2017), 25–36.

[4] Ziyi Bao, Kristijonas Čyras, and Francesca Toni. 2017. ABAplus: Attack Reversal in Abstract and Structured Argumentation with Preferences. In *PRIMA 2017:*

Principles and Practice of Multi-Agent Systems - 20th International Conference (Nice), Bo An, Ana L. C. Bazzan, João Leite, Serena Villata, and Leendert van der Torre (Eds.). Springer, 420–437. https://doi.org/10.1007/978-3-319-69131-2_25

[5] Trevor J M Bench-Capon, Katie Atkinson, and Peter McBurney. 2012. Using argumentation to model agent decision making in economic experiments. *Autonomous Agents and Multi-Agent Systems* 25, 1 (2012), 183–208. https://doi.org/10.1007/s10458-011-9173-6

[6] Philippe Besnard, Alejandro Javier García, Anthony Hunter, Sanjay Modgil, Henry Prakken, Guillermo Ricardo Simari, and Francesca Toni. 2014. Introduction to Structured Argumentation. *Argument & Computation* 5, 1 (2014), 1–4. https://doi.org/10.1080/19462166.2013.869764

[7] Andrei Bondarenko, Phan Minh Dung, Robert Kowalski, and Francesca Toni. 1997. An Abstract, Argumentation-Theoretic Approach to Default Reasoning. *Artificial Intelligence* 93, 97 (1997), 63–101. https://doi.org/10.1016/S0004-3702(97)00015-5

[8] Kristijonas Čyras and Francesca Toni. 2016. ABA+: Assumption-Based Argumentation with Preferences. In *Principles of Knowledge Representation and Reasoning, 15th International Conference*, Chitta Baral, James P Delgrande, and Frank Wolter (Eds.). AAAI Press, Cape Town, 553–556.

[9] Yannis Dimopoulos, Bernhard Nebel, and Francesca Toni. 2002. On The Computational Complexity of Assumption-Based Argumentation for Default Reasoning. *Artificial Intelligence* 141, 1-2 (2002), 57–78. https://doi.org/10.1016/S0004-3702(02)00245-X

[10] Phan Minh Dung, Phan Minh Thang, and Francesca Toni. 2008. Towards Argumentation-Based Contract Negotiation. In *Computational Models of Argument (Frontiers in Artificial Intelligence and Applications)*, Philippe Besnard, Sylvie Doutre, and Anthony Hunter (Eds.), Vol. 172. IOS Press, Toulouse, 134–146.

[11] John Fox, Liz Black, David Glasspool, Sanjay Modgil, Ayelet Oettinger, Vivek Patkar, and Matt Williams. 2006. Towards a general model for argumentation services. In *Argumentation for Consumers of Healthcare, Papers from the 2006 AAAI Spring Symposium*. AAAI, Stanford, 52–57. https://doi.org/10.1.1.148.5967

[12] Paolo Fraccaro, Mercedes Arguello Casteleiro, John Ainsworth, and Iain Buchan. 2015. Adoption of Clinical Decision Support in Multimorbidity: A Systematic Review. *JMIR Medical Informatics* 3, 1 (2015), e4. https://doi.org/10.2196/medinform.3503

[13] Audrey Grace, Carolanne Mahony, John O'donoghue, Tony Heffernan, David Molony, and Thomas Carroll. 2013. A vision for enhancing multimorbid care using clinical decision support systems. *Studies in Health Technology and Informatics* 192, 1-2 (2013), 1117. https://doi.org/10.3233/978-1-61499-289-9-1117

[14] María Adela Grando, David Glasspool, and Aziz A. Boxwala. 2012. Argumentation logic for the flexible enactment of goal-based medical guidelines. *Journal of Biomedical Informatics* 45, 5 (2012), 938–949. https://doi.org/10.1016/j.jbi.2012.03.005

[15] Anthony Hunter and Matthew Williams. 2012. Aggregating evidence about the positive and negative effects of treatments. *Artificial Intelligence in Medicine* 56, 3 (2012), 173–190. https://doi.org/10.1016/j.artmed.2012.09.004

[16] Souhila Kaci and Namrata Patel. 2014. A postulate-based analysis of comparative preference statements. *Journal of Applied Logic* 12, 4 (2014), 501–521. https://doi.org/10.1016/j.jal.2014.07.004

[17] Antonis C Kakas and Pavlos Moraitis. 2003. Argumentation Based Decision Making for Autonomous Agents. In *2nd International Joint Conference on Autonomous Agents & Multiagent Systems*. ACM Press, Melbourne, 883–890. https://doi.org/10.1145/860575.860717

[18] Nadin Kokciyan, Isabel Sassoon, Anthony Young, Martin Chapman, Talya Porat, Mark Ashworth, Vasa Curcin, Sanjay Modgil, Simon Parsons, and Elizabeth Sklar. 2018. Towards an Argumentation System for Supporting Patients in Self-Managing their Chronic Conditions. In *Joint Workshop on Health Intelligence (W3PHIAI)*. New Orleans, Louisiana.

[19] Tuomo Lehtonen, Johannes Peter Wallner, and Matti Järvisalo. 2019. Reasoning over Assumption-Based Argumentation Frameworks via Direct Answer Set Programming Encodings. In *33rd AAAI Conference on Artificial Intelligence*. AAAI Press, Honolulu, Hawaii.

[20] Giorgio Leonardi, Alessio Bottrighi, Gabriele Galliani, Paolo Terenziani, Antonio Messina, and Francesco Della Corte. 2012. Exceptions handling within GLARE clinical guideline framework.. In *AMIA Annual Symposium Proceedings*, Vol. 2012. 512–21.

[21] Luca Longo. 2016. Argumentation for Knowledge Representation, Conflict Resolution, Defeasible Inference and Its Integration with Machine Learning. In *Machine Learning for Health Informatics - State-of-the-Art and Future Challenges*, Andreas Holzinger (Ed.). Vol. 9605. Springer, 183–208. https://doi.org/10.1007/978-3-319-50478-0_9

[22] Sanjay Modgil. 2009. Reasoning About Preferences in Argumentation Frameworks. *Artificial Intelligence* 173, 9-10 (2009), 901–934. https://doi.org/10.1016/j.artint.2009.02.001

[23] Bernard Moulin, Hengameh Irandoust, Micheline Bélanger, and G Desbordes. 2002. Explanation and argumentation capabilities: Towards the creation of more persuasive agents. *Artificial Intelligence Review* 17, 3 (2002), 169–222. https://doi.org/10.1023/A:1015023512975

[24] Jann Muller and Anthony Hunter. 2012. An Argumentation-Based Approach for Decision Making. In *2012 IEEE 24th International Conference on Tools with Artificial Intelligence*. IEEE, Washington, DC, 564–571. https://doi.org/10.1109/ICTAI.2012.82

[25] Christiane Muth, Marjan van den Akker, Jeanet W. Blom, Christian D. Mallen, Justine Rochon, François G. Schellevis, Annette Becker, Martin Beyer, Jochen Gensichen, Hanna Kirchner, Rafael Perera, Alexandra Prados-Torres, Martin Scherer, Ulrich Thiem, Hendrik van den Bussche, and Paul P. Glasziou. 2014. The Ariadne principles: How to handle multimorbidity in primary care consultations. *BMC Medicine* 12, 1 (2014), 1–11. https://doi.org/10.1186/s12916-014-0223-1

[26] Tiago Oliveira, Jérémie Dauphin, Ken Satoh, Shusaku Tsumoto, and Paulo Novais. 2018. Argumentation with Goals for Clinical Decision Support in Multimorbidity. In *17th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, Stockholm, 2031–2033.

[27] Simon Parsons, Carles Sierra, and Nick Jennings. 1998. Agents that reason and negotiate by arguing. *Journal of Logic and Computation* 8, 3 (1998), 261–292. https://doi.org/10.1093/logcom/8.3.261

[28] Mor Peleg. 2013. Computer-interpretable clinical guidelines: A methodological review. *Journal of Biomedical Informatics* 46, 4 (2013), 744–763. https://doi.org/10.1016/j.jbi.2013.06.009

[29] Malik Al Qassas, Daniela Fogli, Massimiliano Giacomin, and Giovanni Guida. 2016. ArgMed: A Support System for Medical Decision Making Based on the Analysis of Clinical Discussions. In *Real-World Decision Support Systems: Case Studies*, Jason Papathanasiou, Nikolaos Ploskas, and Isabelle Linden (Eds.). Springer, 15–41. https://doi.org/10.1007/978-3-319-43916-7_2

[30] Iyad Rahwan and Guillermo Ricardo Simari. 2009. *Argumentation in Artificial Intelligence*. Springer. https://doi.org/10.1007/978-0-387-98197-0

[31] David Riaño and Wilfrido Ortega. 2017. Computer technologies to integrate medical treatments to manage multimorbidity. *Journal of Biomedical Informatics* 75, September (2017), 1–13. https://doi.org/10.1016/j.jbi.2017.09.009

[32] Lucia Sacchi, Stefania Rubrichi, Carla Rognoni, Silvia Panzarasa, Enea Parimbelli, Andrea Mazzanti, Carlo Napolitano, Silvia G. Priori, and Silvana Quaglini. 2015. From decision to shared-decision: Introducing patients' preferences into clinical decision analysis. *Artificial Intelligence in Medicine* 65, 1 (2015), 19–28. https://doi.org/10.1016/j.artmed.2014.10.004

[33] Erez Shalom, Yuval Shahar, and Eitan Lunenfeld. 2016. An architecture for a continuous, user-driven, and data-driven application of clinical guidelines and its evaluation. *Journal of Biomedical Informatics* 59, November (2016), 130–148. https://doi.org/10.1016/j.jbi.2015.11.006

[34] Matteo Spiotta, Paolo Terenziani, and Daniele Theseider Dupré. 2017. Temporal Conformance Analysis and Explanation of Clinical Guidelines Execution: An Answer Set Programming Approach. *IEEE Transactions on Knowledge and Data Engineering* 29, 11 (2017), 2567–2580. https://doi.org/10.1109/TKDE.2017.2734084

[35] Pancho Tolchinsky, Ulises Cortés, Sanjay Modgil, Francisco Caballero, and Antonio López-Navidad. 2006. Increasing Human-Organ Transplant Availability: Argumentation-Based Agent Deliberation. *IEEE Intelligent Systems* 21, 6 (2006), 30–37. https://doi.org/10.1109/MIS.2006.116

[36] Rosy Tsopra, Jean Baptiste Lamy, and Karima Sedki. 2018. Using preference learning for detecting inconsistencies in clinical practice guidelines: Methods and application to antibiotherapy. *Artificial Intelligence in Medicine* 89, February (2018), 24–33. https://doi.org/10.1016/j.artmed.2018.04.013

[37] Kristijonas Čyras and Tiago Oliveira. 2018. Argumentation for Reasoning with Conflicting Clinical Guidelines and Preferences. In *Principles of Knowledge Representation and Reasoning, 16th International Conference* (Tempe, AZ), Michael Thielscher, Francesca Toni, and Frank Wolter (Eds.). AAAI Press, 631–632.

[38] Neeltje P. C. A. Vermunt, Mirjam Harmsen, Gert P. Westert, Marcel G. M. Olde Rikkert, and Marjan J. Faber. 2017. Collaborative goal setting with elderly patients with chronic disease or multimorbidity: a systematic review. *BMC Geriatrics* 17, 1 (2017), 167. https://doi.org/10.1186/s12877-017-0534-0

[39] Toshiko Wakaki. 2014. Assumption-Based Argumentation Equipped with Preferences. In *Principles and Practice of Multi-Agent Systems - 17th International Conference*, Hoa Khanh Dam, Jeremy V Pitt, Yang Xu, Guido Governatori, and Takayuki Ito (Eds.), Vol. 8861. Springer, Gold Coast, 116–132. https://doi.org/10.1007/978-3-319-13191-7_10

[40] Douglas Walton. 1996. *Argumentation Schemes for Presumptive Reasoning*. L. Erlbaum Associates.

[41] Szymon Wilk, Martin Michalowski, Wojtek Michalowski, Daniela Rosu, Marc Carrier, and Mounira Kezadri-Hamiaz. 2017. Comprehensive mitigation framework for concurrent application of multiple clinical practice guidelines. *Journal of Biomedical Informatics* 66 (2017), 52–71. https://doi.org/10.1016/j.jbi.2016.12.002

[42] Veruska Zamborlini, Marcos da Silveira, Cedric Pruski, Annette ten Teije, Edwin Geleijn, Marike van der Leeden, Martijn Stuiver, and Frank van Harmelen. 2017. Analyzing interactions on combining multiple clinical guidelines. *Artificial Intelligence in Medicine* 81 (2017), 78–93. https://doi.org/10.1016/j.artmed.2017.03.012

[43] Zhiwei Zeng, Xiuyi Fan, Chunyan Miao, Cyril Leung, Chin Jing Jih, and Ong Yew Soon. 2018. Context-based and Explainable Decision Making with Argumentation. In *17th International Conference on Autonomous Agents and MultiAgent Systems*. IFAAMAS, Stockholm, 1114–1122.