Assumption-based argumentation with preferences and goals for patient-centric reasoning with interacting clinical guidelines

Kristijonas Čyras [a,*], Tiago Oliveira [b,**], Amin Karamlou [c,***] and Francesca Toni [c]

[a] *Department of Computing, Imperial College London, United Kingdom*
E-mail: kcyras@gmail.com; ORCID: https://orcid.org/0000-0002-4353-8121

[b] *National Institute of Informatics, Japan*
E-mail: tiago.jose.martins.oliveira@gmail.com

[c] *Department of Computing, Imperial College London, United Kingdom*
E-mails: karamlouamin@gmail.com, f.toni@imperial.ac.uk

Abstract. A paramount, yet unresolved issue in personalised medicine is that of automated reasoning with clinical guidelines in multimorbidity settings. This entails enabling machines to use computerised generic clinical guideline recommendations and patient-specific information to yield patient-tailored recommendations where interactions arising due to multimorbidities are resolved. This problem is further complicated by patient management desiderata, in particular the need to account for patient-centric goals as well as preferences of various parties involved. We propose to solve this problem of automated reasoning with interacting guideline recommendations in the context of a given patient by means of computational argumentation. In particular, we advance a structured argumentation formalism ABA^+G (short for Assumption-Based Argumentation with Preferences (ABA^+) and Goals) for integrating and reasoning with information about recommendations, interactions, patient's state, preferences and prioritised goals. ABA^+G combines assumption-based reasoning with preferences and goal-driven selection among reasoning outcomes. Specifically, we assume defeasible applicability of guideline recommendations with the general goal of patient well-being, resolve interactions (conflicts and otherwise undesirable situations) among recommendations based on the state and preferences of the patient, and employ patient-centered goals to suggest interaction-resolving, goal-importance maximising and preference-adhering recommendations. We use a well-established Transition-based Medical Recommendation model for representing guideline recommendations and identifying interactions thereof, and map the components in question, together with the given patient's state, prioritised goals, and preferences over actions, to ABA^+G for automated reasoning. In this, we follow principles of patient management and establish corresponding theoretical properties as well as illustrate our approach in realistic personalised clinical reasoning scenaria.

Keywords: Automated medical reasoning, structured argumentation, principles of patient management

---

*Corresponding author: Kristijonas Čyras, Current affiliation and address: Ericsson, Torshamnsgatan 23, 164 83, Stockholm, Sweden. Tel.: +46724666586; E-mail: kcyras@gmail.com.

**Tiago Oliveira has pursued the research presented in this article while at the National Institute of Informatics, but is no longer part of this institution.

***Kristijonas Čyras and Amin Karamlou have pursued research presented in this article while at Imperial College London and at the time of writing this article are alumni of this institution.

1. Introduction

In the context of medical reasoning, patient management involves careful consideration of the patient's condition and applicable treatments which should lead to a desired state. Clinical guidelines (such as [44]) are used as the textbook source offering best practice recommendations in general patient management. These documents are by-and-large designed to target single health conditions, leading to issues in the presence of multiple health conditions (multimorbidities). Indeed, in such situations, clinical guidelines should be combined, hence raising the need to consider multiple interactions that impact the evolution of a patient [41,49]. These interactions may render suggested recommendations inapplicable, conflicting, overlapping and so forth. Thus, multimorbidities create obstacles to clinicians in the application of clinical guideline recommendations. In this context, knowledge representation methods from AI may offer mechanisms to ease these obstacles.

Easing the application of clinical guidelines is the objective of the *Transition-based Medical Recommendation model* (TMR) [108,109], a state-of-the-art formalism [82] for representing computerised clinical guideline recommendations. TMR components and relations reflect knowledge and occurrences typical of multimorbidity situations: the basic components are clinical care actions and their respective effects on the patient's physical properties; the relations amount to interactions among those actions and their effects. Given the paramount importance of recommendation interactions, TMR provides a mechanism to identify various types of interactions, such as contradiction, repetition and alternative. Therefore, TMR is a comprehensive model for clinical guideline recommendations and situations spawning from their application. However, TMR does not provide *reasoning* mechanisms to resolve the interactions automatically and thence select recommendations for specific patients.

Reasoning is also limited in several other proposed formalisms for clinical guideline representation, particularly when conflicts come into play [41,75,82]. (A notable exception is the recent CONSULT project [19,20,59,107], which we discuss in Section 7.2.) Additionally, the representations afforded by such formalisms rarely take into account the context of the patient, namely patient-specific conditions, patient-centric goals, and preferences from the various parties involved [75,83,99]. Indeed, integrating all these elements is no easy task. The *Ariadne principles* [72] attempt to take into account all these elements and provide a conceptual structure for patient management in the context of multimorbidities, stressing the importance of interaction assessment, individual management and patient's and/or clinician's goals and preferences. Inspired by these Ariadne principles, in this work we propose a formal framework using a TMR-based and *argumentation*-enabled approach to reason with interacting clinical guideline recommendations in the context of specific patients, taking into account their state, goals and preferences.

Argumentation is fit for this task as it allows for reasoning with uncertain and conflicting information. Argumentation models reasoning of autonomous agents in multi-agent systems in a way that emulates human reasoning, see e.g. [10,55,74,80]. It has been widely applied to support medical reasoning, see e.g. [29,40,52,63,73,94]. The interest in argumentation from a medical domain perspective is related to the ability of argumentation to allow "for important conflicts to be highlighted and analysed and unimportant conflicts to be suppressed" [6]. We employ structured argumentation (see e.g. [80, Part II] and [11] for overviews) in the form of *Assumption-Based Argumentation with Preferences* (ABA$^+$) [15, 27,33] to automate patient-centric reasoning based on conflicting guideline recommendations, goals, and preferences.

The choice of ABA$^+$ is motivated by several of its characteristics. On the one hand, the nature of knowledge representation and reasoning in ABA$^+$ suits the task of reasoning with interacting clinical

guidelines well. Indeed, the rule-based specification of ABA$^+$ frameworks allows for a natural representation of TMR concepts, particularly recommendations, which are essentially of the form "assuming you follow recommendation R, perform action A, which will bring about effect E that affects property P, leading to a change from the initial value v_I to the target value v_T", and can be seen as rules 'if R then A', 'if A then E', 'if P takes v_I and E, then P will take v_T'. Representation in other argumentation formalisms, e.g. in Value-Based Argumentation [8,9,54]. Further, since in the context of multiple applicable yet interacting clinical guidelines one needs to make a defensible choice as to which ones to follow, credulous reasoning, particularly in terms of preferred extensions, is very adequate. Such reasoning using extension-based semantics is naturally supported in ABA$^+$ but not in Defeasible Logic Programming (DeLP) [43] or Carneades [45,47]. ABA$^+$ also offers a built-in reasoning mechanism to deal with preferences which, differently from other structured argumentation formalisms, e.g. ASPIC$^+$ [67,68,76], force attacks to be reversed in specific cases, all the while preserving conflict-freeness of sets of assumptions and ensuring desirable properties thereof. This allows for a simple representation of, and reasoning with, preferences among recommendations in the presence of interactions, as well as satisfaction of the Ariadne principles.

On the other hand, we are strongly driven by practical concerns of deployment of our envisaged argumentation-assisted clinical decision support system. To this end ABA$^+$ is a particularly suitable choice. For one, ABA$^+$ has some known complexity results, first established for the underlying ABA formalism [15] in [34] and recently for its extension with preferences (ABA$^+$) in [60]. Very importantly, ABA$^+$ is equipped with working implementations, for instance the stand-alone[1] and web[2] applications as described in [7] and a stand-alone development[3] built on [56]. These make it easy to implement ABA$^+$G, connect it to TMR (via an implementation[4] of [109] and its programming interface TMRweb [20]) and thus lay grounds for the decision support system in question.

We use ABA$^+$ to reason with the TMR representations of recommendations and interactions via rules and arguable elements (i.e. assumptions representing applicability of recommendations) from which arguments (as deductions) are constructed. We integrate patient-specific information as well as preferences over actions (effectively, over recommendations) alongside TMR representation in ABA$^+$. We use extension-based semantics for reasoning, thus providing an assumption-driven method by which the applicability of recommendations is argued for or against in light of a patient's condition. This ensures that all the interactions amongst the suggested recommendations have been resolved. To incorporate treatment goals, we augment ABA$^+$ to form ABA$^+$G by introducing a goal-driven reasoning mechanism to select the best interaction-free (sets of) recommendations based on the importance of patient-centric goals. These knowledge representation, reasoning as well as conflict and preference handling mechanisms used in our approach allow us to meet the Ariadne principles. We illustrate our approach to patient-centric reasoning with interacting recommendations, goals and preferences using a TMR-based case study and show arguably desirable outcomes.

We summarise the main contributions of this paper as follows:

- We enable automated reasoning with interacting clinical guidelines represented in the Transition-based Medical Recommendation model (TMR), by mapping recommendations and interactions (of types contradiction, repetition, alternative, and repairable) to a structured-argumentation formalism, ABA$^+$;

---

[1]github.com/kcyras/ABAplus
[2]www-abaplus.doc.ic.ac.uk
[3]github.com/kcyras/ABAPlusG
[4]github.com/Data2Semantics/guidelines/

- We embed patient's conditions and preferences in ABA$^+$ for assumption-based reasoning with conflicting recommendations and patient-specific information;
- We augment ABA$^+$ with prioritised goals for goal-driven patient-centric reasoning with recommendations, to obtain ABA$^+$G;
- We establish some theoretical properties of ABA$^+$G, relating them to the Ariadne principles of patient management;
- We illustrate the reasoning with a realistic set of guideline recommendations in different patient contexts;
- We scrutinise some conceptual and technical choices of our approach and discuss it in relation to argumentative and non-argumentative works in medical reasoning and decision making.

The present work is based on and significantly extends the work in [30] by incorporating additional TMR artefacts, broadening the theoretical exposition of ABA$^+$G and providing an extensive case study illustration. Specifically in terms of TMR, we deal with target values of the properties affected by recommended actions (see Section 3.1.1) and several types of interactions (see Section 3.1.2). As regards ABA$^+$G, we additionally model non-applicability of recommendations and the logic of repairable interactions (see Section 4.3), and slightly generalise the theoretical results regarding the desirable properties of dealing with interacting recommendations (see Section 4.3.3). The case study illustration (see Section 5) is completely new and provides a detailed exemplification of all these aspects.

Currently, an end-to-end proof-of-concept system encompassing electronic health record (EHR) information about patients, TMR via its implementation TMRweb, and ABA$^+$G to provide decision support to clinicians is under development within the ROAD2H project.[5] In this paper we provide the theoretical framework for both ABA$^+$G and its implementation[6] which is compatible with a wrapper interface that integrates TMRweb, EHR hooks and other relevant functionalities (such as for preference elicitation). The specification of algorithms and other engineering details pertaining to this implementation of ABA$^+$G is beyond the scope of this paper and is left for better suited future publications describing the overall decision support system.

We structure this paper as follows. In Section 2 we consider desiderata for our approach in terms of patient management principles from medical literature. We then describe, in Section 3, 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 some design choices as well as limitations of our approach in Section 6. In Section 7 we place our work in the context of several related works. We end in Section 8 with conclusions and a summary of future work directions.

2. Principles of patient management

In this work we consider the medical reasoning aspect of patient management in a multimorbidity setting. Various works acknowledge several principles of patient management [41,49,75,83,99], but their respective analyses are neither systematic nor provide the necessary level of detail. In contrast, [72] stands out with a comprehensive enumeration and description of patient management principles, therein called **Ariadne principles**. Our interpretation of them is as follows.

---

1. **Interaction assessment**: recommendation interactions and their 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-interacting recommendations for the given patient.

Rather than providing specific methods to handle conflicts stemming from clinical guideline recommendations, the Ariadne principles point out which aspects should be considered in medical reasoning involving multimorbidities and patient context. As for treatment goals, it is stated that information about the effect of treatments on general goals such as increasing life expectancy or quality of life are often unavailable. Instead, restricting treatment goals to tangible effects brought about (or not) by treatments, such as symptom relief, disease prevention, and avoidance of unwanted outcomes, seems to be more effective in this situation. Additionally, the Ariadne principles establish that patient and physician should discuss preferences over actions and priorities over treatment goals, which should be taken into account when devising a treatment plan for the patient.

Reasoning with clinical guidelines in the context of multimorbidities involves aggregation of discordant guideline recommendations and respective interactions. While TMR provides an expressive representation template for this information, it does not enable the above-mentioned aggregation for reasoning to produce patient-specific solutions in a multimorbidity setting. Thus, adhering to the Ariadne principles, even when using TMR for representation, calls for establishing foundations for reasoning in the context of a patient. We answer this call in this paper by situating the TMR model and patient context for reasoning with in ABA$^+$G.

3. Problem setting

We here describe the problem of reasoning with interacting clinical guideline recommendations in the context of a patient. We first review the TMR model and interactions among recommendations. We then discuss the context of a patient.

Alongside theoretical developments we are concerned with an end-to-end implemented system for reasoning with interacting guideline recommendations. We thus provide details on TMR following [108] but focus on the core features that are already largely implemented and present in [109] and TMRweb, and that will be handled by ABA$^+$G. In what follows we may detail which features of the latest as yet unimplemented theoretical development [108] of TMR we are not making use of (indicated with *).

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 [108], we assume that

Fig. 1. TMR representation schema instantiated with recommendations R_1 and R_2 [108, p. 83, Fig. 2]. (Figure kindly provided by the authors of [108].)

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. (Here, the recommendation concerning NSAID[7] is taken from a Diabetes guideline, and the recommendation concerning Aspirin is taken from an Osteoarthritis guideline.) It consists of the following components.[8]

1. **Name**, e.g. R_1, R_2, at the top of a rounded box.
   (We write R_k instead of Rk.) We make a tacit assumption that recommendation names are unique and distinct from all symbols appearing in the other components. Henceforth, we refer to a recommendation by its name.
2. A unique associated **action** A, e.g. *Adm. Aspirin*, *Adm. NSAID* (where *Adm.* stands for Administer).
3. **Deontic strength**, which we denote by δ, is indicated by a thick labelled arrow and "reflects a degree of obligatoriness expected for that recommendation" [108, p. 82]. It takes values in $[-1, 1]$: if $\delta \geqslant 0$, then the recommendation R with deontic strength δ recommends performing the action; if $\delta < 0$, then R recommends avoiding the action. To discretise δ, we use two qualitative landmarks *should* and *should not*, corresponding to values 0.5 and -0.5, respectively, as available in the current TMR implementation. For illustration, the deontic strengths of R_1 and R_2 in Fig. 1 are $\delta_1 = 0.5 = should$ and $\delta_2 = -0.5 = should\ not$, respectively.
4. **Contributions** of the recommendation to the overall goals in the context of a guideline. A recommendation can have multiple contributions, each carrying an *identifier*, e.g. $C1.1$, $C2.1$, indicated below the recommendation name. A contribution consists of the following components.

   (i) **Property** affected by the action, e.g. *Blood Coagulation*, *Gastrointestinal Bleeding*.
   (ii) **Effect** of the action on the property, e.g. *decrease*, *increase*.
   (iii–iv) **Initial** and **target values** of the property that the action affects. For instance, *Adm. NSAID* leads to a *decrease* in *Blood Coagulation* from the initial value *normal* to the target value *low*. Otherwise, ? represents an *indeterminate* value.[9]

---

[7]Non-steroidal anti-inflammatory drug. NSAIDs are medicines that are widely used to relieve pain, reduce inflammation, and bring down a high temperature, see e.g. www.nhs.uk/conditions/nsaids/.

[8] The original description of recommendations, with components as functions/relations, more suitable for implementation efforts, is long and unnecessary for the purposes of this paper. Instead, we give an intermediate representation which carries the necessary aspects required in this work, following the alternative formal description (and visualisation) in [108] of TMR instances, faithful to the original but omitting certain aspects (as indicated below).

[9]In practice, indeterminate values do not appear at all, because concrete raw values appearing in the patient's EHR are processed by a parser to instantiate an 'intermediate' patient's record with the qualitative values as they appear in TMRweb. We will thus henceforth instantiate any indeterminate values with specific qualitative values, without any loss of generality.

* In this paper we *will not* make use of, but mention for completeness, two quantitative values associated with the effect of the contribution: *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. We will also not make use of the overall value of the contribution, in the range of $[-1, 1]$ (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.

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

1. name R,
2. action A,
3. deontic strength δ,
4. a set of contributions $\mathcal{C} = \{C^1, \ldots, C^n\}$, for $n \geqslant 1$, where a contribution is a tuple (P, E, v_I, v_T) with

   (i) property P affected,
   (ii) effect E on the property,
   (iii) initial value v_I of the property that the action's effect applies to,
   (iv) target value v_T of the property expected after the effect applies.

Whenever $|\mathcal{C}| = 1$, we may abuse the notation and write $(R, A, \delta, (P, E, v_I, v_T))$ for a recommendation.

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

Example 3.1. $R_1 = (R_1, Adm. NSAID, should, (Blood Coagulation, decrease, normal, low))$ and $R_2 = (R_2, Adm. Aspirin, should not, (Gastrointestinal Bleeding, increase, normal, high))$ are illustrated in Fig. 1 (we instantiated the indeterminate ? with specific values *normal* and *high* in R_2, cf. footnote 9). We thus have $\mathbb{R} = \{R_1, R_2\}$.

3.1.2. Interactions

Using TMR, one can identify *interactions* among recommendations [108,109]. Intuitively, interactions record various relationships between different recommendations. In particular:

- *Contradiction* in case a particular recommendation urges avoiding the action suggested by another recommendation.
- *Repetition* in case recommendations suggest taking or avoiding the same action.
- *Alternative* in case recommendations concern different actions having the same or similar consequences.
- *Repairable* in case the consequences of following one recommendation revert the (negative) consequences of following another recommendation.

Interactions and their identification are formally defined in [108,109], but those details are not important for the purposes of this paper. We treat interactions of various types as outputs of (the implementation of) TMR for argumentation to reason with. While several types of interactions can be identified in principle [108], the existing implementation of TMR affords identification of, specifically, Contradiction, Repetition, Alternative and Repairable types of interactions. These are the types of interactions we focus on in this paper and show how they can be naturally resolved by means of argumentation.

Formally, we define:

Definition 3.2. An **interaction** between recommendations R_i, $R_j \in \mathbb{R}$ is a tuple (R_i, R_j, t), where $t \in \mathcal{T} = \{Contr, Repet, Alt, Repair\}$ is the *type* of the interaction. *Contr*, *Repet*, *Alt* and *Repair* stand for Contradiction, Repetition, Alternative and Repairable, respectively.

From now on, $\mathbb{I}$ denotes the set of all interactions (given $\mathbb{R}$).

Example 3.2. The recommendations R_1 and R_2 from Example 3.1 are in a Contradiction interaction, as they recommend opposite actions.[10] We thus assume that $(R_1, R_2, Contr) \in \mathbb{I}$.

Remark 1. $\mathbb{I}$ is symmetric in the first two components for $t \in \mathcal{T} \setminus \{Repair\}$ in the sense that for $t \in \{Contr, Repet, Alt\}$, both (R_i, R_j, t) and (R_j, R_i, t) express the same interaction, namely that R_i and R_j are in, respectively, Contradiction, Repetition or Alternative interaction. Accordingly, TMRweb yields only one of the interactions in such cases. However, Repairable interactions are not symmetric in the same sense, and the TMRweb output $(R_i, R_j, Repair)$ means that R_j 'repairs' R_i, but not vice versa.

When reasoning with interacting clinical guideline recommendations, the goal is to resolve the interactions to be able to follow the recommendations. In particular, Contradiction, Repetition and Alternative interactions are the kind that a clinician aims to avoid having among the recommendations they intend to follow. In other words, no two recommendations R and R' in interaction of type Contradiction, Repetition or Alternative should be mutually followed. On the other hand, an interaction of type Repairable tells a clinician that potential problems arising by following one recommendation can be resolved by following another recommendation that 'repairs' the first one.

The above interpretation of interactions gives rise to the following notions of *interaction-free* and *interaction-resolving* sets of recommendations.

Definition 3.3. Let $\mathbb{R}' \subseteq \mathbb{R}$ be a set of recommendations.

- $\mathbb{R}'$ is **interaction-free** iff there is no interaction $(R_i, R_j, t) \in \mathbb{I}$ of type $t \in \{Contr, Repet, Alt\}$ with $R_i, R_j \in \mathbb{R}'$.
- $\mathbb{R}'$ is **interaction-resolving** iff $\mathbb{R}'$ is interaction-free, and whenever $R_i \in \mathbb{R}'$ and there is a Repairable interaction $(R_i, R_j, Repair) \in \mathbb{I}$, then for at least one $(R_i, R_j, Repair) \in \mathbb{I}$ it holds that $R_j \in \mathbb{R}'$.

Intuitively, interaction-free sets of recommendations consist of recommendations that are safe to follow without the risk of performing a) incompatible (in the case of contradictions), or b) superfluous (in the case of alternatives and repetitions) actions. In addition, interaction-resolving sets of recommendations aim to avoid the risk of performing c) insufficient actions (in the case of repairability). For a recommendation that is repairable, one repair suffices to resolve the interaction, but there may in principle be multiple repairs in an interaction-resolving set of recommendations.

Example 3.3. The set $\mathbb{R} = \{R_1, R_2\}$ from Example 3.1 is not interaction-free, for $(R_1, R_2, Contr) \in \mathbb{I}$, as in Example 3.2. Clearly, $\{R_1\}$ and $\{R_2\}$ themselves are interaction-free and interaction-resolving.

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

---

[10]Note that a hierarchy of actions is assumed in [108, p. 79] to obtain interactions. For instance, the action to administer NSAID subsumes both actions to administer Aspirin and Ibuprofen. This hierarchy is used when specifying actions in TMRweb, but is not important for our purposes.

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. That is, in order to apply recommendations, one needs to consider specific patient *conditions* and the initial values of the effects that actions have on properties. For instance, a patient can have conditions *normalBlood Coagulation* or *normalGastrointestinal Bleeding* (here and henceforth we concatenate the property with its initial value to represent a patient's condition).

Example 3.4. Consider $\mathbb{R} = \{R_1, R_2\}$ and $\mathbb{I} = \{(R_1, R_2, Contr)\}$ as in Examples 3.1 and 3.2. Intuitively, for a patient with normal blood coagulation (*normalBlood Coagulation*), NSAID – e.g. Aspirin – should be administered. If, however, the patient shows gastrointestinal bleeding (say *normalGastrointestinal 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. [83]). To resolve the conflict in Example 3.4, 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 3.5. Continuing Example 3.4, 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.

In general, preference information of various parties often needs to be taken into account to deliver the best care, see e.g. [75,83]. 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. Thus, when using argumentation frameworks to reason with guidelines in Section 4, patient conditions will come as information additional to TMR instances. 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 3.4. 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 $\mathcal{G}$,
- preferences $\leqslant$ over actions,
- priorities $\preccurlyeq$ over goals.[11]

---

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

Preferences over actions and priorities over goals can come from various sources, such as the patient, the patient's family, the clinician or the clinic, and may involve various considerations, such as the cost, availability or quality of evidence regarding actions and importance of goals. For simplicity, in this paper we often ascribe preferences over actions to the patient and priorities over goals to the clinician, without qualifying the underlying considerations. 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, or possibly an empty set) over the effects on the patient's state. We make this precise in Section 4.3.

Example 3.6. Continuing from Example 3.5, by concatenating properties with values or effects the context of the patient can be given by

- $\mathcal{S} = \{normalBlood\ Coagulation,\ normalGastrointestinal\ Bleeding\}$,
- $\mathcal{G} = \{decreaseBlood\ Coagulation,\ \mathsf{not}\ increaseGastrointestinal\ Bleeding\}$,[12]
- $R_2 < R_1$,[13]
- *decreaseBlood Coagulation* $\prec$ not *increaseGastrointestinal 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 2nd and the 3rd Ariadne principles put forward in Section 2.

4. ABA$^+$G for reasoning

We will use guideline recommendations, their interactions and contexts to construct argumentation frameworks for an agent to reason and resolve interactions among recommendations, given patient-specific conditions, patient-centric goals and various preferences. Specifically, we will use ABA$^+$ frameworks, which we review in Section 4.1, for assumption-based reasoning with guidelines and patient's preferences over recommendations. We will then, in Section 4.2, augment ABA$^+$ to ABA$^+$G for goal-driven reasoning with guidelines and clinician's priorities over goals. We finally describe and formalise patient-centric reasoning with interacting guideline recommendations in ABA$^+$G, and establish its properties that pave the way to meet the Ariadne principles, in Section 4.3.

4.1. ABA$^+$ background

We provide the background for ABA$^+$ following [15,33].
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*;

---

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

[13]As usual, the strict (asymmetric) counterpart $<$ of a preorder $\leqslant$ is given by $\alpha < \beta$ iff $\alpha \leqslant \beta$ and $\beta \not\leqslant \alpha$, for any α and β. The preorder $\leqslant$ is thus given by the reflexive and transitive closure of $<$. We assume this for all preorders in this paper.

- $\mathcal{A} \subseteq \mathcal{L}$ 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 A;
- $\leqslant$ 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 specified otherwise.

Assumptions in ABA⁺ represent arguable information. For instance, assumptions can represent the agent's potential to follow a recommendation. In such a case, preferences in ABA⁺ can represent the relative (patient's) willingness to follow different 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 (unspecified) $R \subseteq \mathcal{R}$.

For $A, B \subseteq \mathcal{A}$, A *<-attacks* B, denoted $A \rightsquigarrow_< B$, [14] iff:

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 against something in B by means of no inferior elements (*normal attack*), b) or B argues against 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 (as in standard ABA [15,27,95]).

We next give notions used to define ABA⁺ semantics in terms of *extensions*, i.e. sets of arguments meeting given requirements.

Let $A \subseteq \mathcal{A}$. The *conclusions of A* amount to 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 that A is *closed* iff $A = Cn(A) \cap \mathcal{A}$, i.e. A contains all its conclusions. We say that $\mathcal{F}$ is *flat* iff every $A \subseteq \mathcal{A}$ is closed. Further:

1. A is *<-conflict-free* iff $A \not\rightsquigarrow_< A$;
2. A *<-defends* $A' \subseteq \mathcal{A}$ iff for all closed $B \subseteq \mathcal{A}$ with $B \rightsquigarrow_< A'$ we have $A \rightsquigarrow_< B$;
3. A is *<-admissible* iff it is closed, *<-conflict-free* and *<-defends* itself.

We consider one particular ABA⁺ semantics, namely *<-preferred extensions*:

4. A set $E \subseteq \mathcal{A}$ of assumptions is a **<-preferred extension** of $\mathcal{F} = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^{-}, \leqslant)$ iff E is $\subseteq$-maximally *<-admissible*.

Note that the above effectively defines non-flat ABA⁺, i.e. generic (non-flat) ABA frameworks with preferences as introduced in [25].

---

[14]The subscript $<$ on $\rightsquigarrow_<$ indicates that $\rightsquigarrow_<$ takes preferences into account in ABA⁺, in contrast to the attack relation $\rightsquigarrow$ in standard ABA [15,27,95], where preferences are absent.

Remark 2. Whenever the preference relation $\leqslant$ on $\mathcal{A}$ is an equivalence – i.e. reflexive, symmetric and transitive – there are no reverse attacks, and preferences do not really play a role in normal attacks either. In other words, any $<$-attack $A \rightsquigarrow_< B$ amounts to there being an argument $A' \vdash \overline{b}$ for some $b \in B$, supported by $A' \subseteq A$, which is the definition of *attack* in ABA [15,27,95]. Similarly, if the preference relation were allowed to be empty, i.e. $\leqslant = \emptyset$, then (ABA$^+$) $<$-attacks would boil down to (ABA) attacks. In other words, ABA$^+$ frameworks with equivalence preference relations $\leqslant$ are semantically equivalent to ABA$^+$ frameworks 'with no preferences'. However, since $\emptyset$ is not reflexive, hence not a preorder, $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \emptyset)$ is not a well defined ABA$^+$ framework. Nonetheless, if we take the preference relation to be the reflexive closure of $\emptyset$, i.e. $\leqslant = RCl(\emptyset) = \{(a, a) \ : \ a \in \mathcal{A}\}$ (which is an equivalence relation), then $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$ is well defined and with an abuse of notation we can call it an ABA$^+$ framework 'with no preferences'. We will make use of this notation in Section 5.

4.2. ABA$^+$G: ABA$^+$ with goals

We extend ABA$^+$ with a mechanism to distinguish among preferred extensions based on goals fulfilled. Goal seeking mechanisms in structured argumentation are introduced in [73] to rank extensions according to the relative priorities over goals fulfilled in the extensions. 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 a rule r in $\mathcal{R}$ with head θ;
- $\preccurlyeq$ is a total preorder on $\mathcal{G}$, denoting **priorities** over goals; for $\theta, \chi \in \mathcal{G}$, $\theta \preccurlyeq \chi$ means that χ is as important as θ. (By convention, if $\mathcal{G} = \emptyset$, then $\preccurlyeq = \emptyset$ too.)

In what follows, $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$ is a fixed but otherwise arbitrary ABA$^+$G framework, unless stated otherwise. Also, slightly abusing the notation, we define a $<$-preferred extension of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$ to be a $<$-preferred extension E of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant)$.

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 $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$. 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 in the conclusions of 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

$$\mathcal{G} \trianglelefteq_{\mathbb{G}} \mathcal{G}' \quad \text{iff} \quad \exists g' \in \mathcal{G}' \setminus \mathcal{G} \quad \text{with } g \preccurlyeq g' \ \forall g \in \mathcal{G} \setminus \mathcal{G}'.$$

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

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

Our 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 fulfil 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. [53] for a comparison of various orderings, but we chose the above in accordance with our interpretation of priorities over goals.

4.3. Representing and reasoning with TMR in ABA$^+$G

We now introduce the representation in ABA$^+$G of TMR instances, interactions and context. We start with an intuitive illustration, then give the formalisation and establish how it meets the Ariadne principles.

4.3.1. Intuition

At a high-level, assumptions will represent (the defeasible potential to follow) recommendations, whereas the corresponding actions and their effects on properties will be modelled via rules, and the deontic strength will determine whether the actions and their consequences are sought after or not, as represented by adding the syntactic not to the heads of rules. While recommendations are assumed to be potentially applicable by default (given an appropriate guideline), they may in principle be *inapplicable* for a given patient, unless the patient presents with a condition, i.e. property, affected by the action associated with the recommendation. This behaviour will be modelled via additional assumptions concerning the possible non-applicability of recommendations. (See also Section 6.2.1 later for a discussion on applicability of specific instances of recommendations.) The context of the patient 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 $\mathbb{R} = \{R_1, R_2\}$, where

- $R_1 = (R_1, \textit{Adm. NSAID}, \textit{should}, (\textit{Blood Coagulation}, \textit{decrease}, \textit{normal}, \textit{low}))$ and
- $R_2 = (R_2, \textit{Adm. Aspirin}, \textit{should not}, (\textit{Gastrointestinal Bleeding}, \textit{increase}, \textit{normal}, \textit{high}))$,

and interactions $\mathbb{I} = \{(R_1, R_2, \textit{Contr})\}$ as in Example 3.4.

First, $R_1, R_2 \in \mathcal{A}$ represent the potential to apply the 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* $\leftarrow R_2$.[15]

The following rules model the effects the actions *Adm. NSAID* and *Adm. Aspirin* bring about and allow to avoid, respectively:

3. *decreaseBlood Coagulation* $\leftarrow$ *Adm. NSAID*;
4. not *increaseGastrointestinal Bleeding* $\leftarrow$ not *Adm. Aspirin*.

Then, the following rules encode that the specific target values of the properties can be expected (to be avoided) given the effects of the actions and the initial values of the properties:

---

[15] As already mentioned in footnote 12, not is purely syntactic.

5. *lowBlood Coagulation ← normalBlood Coagulation, decreaseBlood Coagulation*;
6. not *highGastrointestinal Bleeding ← normalGastrointestinal Bleeding*,
                              not *increaseGastrointestinal Bleeding*.

Now, the additional assumptions $\mathsf{inapp}(R_1)$, $\mathsf{inapp}(R_2) \in \mathcal{A}$ represent the potential non-applicability of recommendations, expressed via the following rules:

7. $\overline{R_1} \leftarrow \mathsf{inapp}(R_1)$;
8. $\overline{R_2} \leftarrow \mathsf{inapp}(R_2)$.

Here, $\overline{R_1}$ and $\overline{R_2}$ are the contraries of R_1 and R_2, respectively.

Arguments against the presumed non-applicability of recommendations will be available whenever the presence of the potentially affected properties can be argued for, allowed by the following rules:

9. $\overline{\mathsf{inapp}(R_1)} \leftarrow$ *normalBlood Coagulation*;
10. $\overline{\mathsf{inapp}(R_2)} \leftarrow$ *normalGastrointestinal Bleeding*.

Now, R_1 and R_2 are in contradiction with *Adm. NSAID* and *Adm. Aspirin*, recommended positively and negatively, respectively. Thus, each can be argued against on the basis of the other, in the presence of the interaction. Therefore, we have:

11. $\overline{R_2} \leftarrow R_1$;
12. $\overline{R_1} \leftarrow R_2$.

Dealing with Repetition and Alternative interactions is similar to dealing with Contradiction interactions as suggested above. The intuition is that repetitive or alternative actions are superfluous, and could possibly lead to adverse effects, whence they should not be taken in tandem. That is, recommendations suggesting repetitive or alternative actions will be mutually conflicting in ABA$^+$G. This is in accordance with the desirable reading of interactions as in Section 3 and in [108, p. 91].

Repairable interactions, on the other hand, are more nuanced. Intuitively, following a recommendation R_i that is repairable by another recommendation R_j seems to necessitate following R_j too. However, in case R_i is repairable by multiple recommendations that are potentially alternatives to one another, it should arguably suffice to follow only one of them. That is, if some R_j that repairs R_i is accepted (i.e. appears in a $<$-preferred extension), then R_i should be considered 'repaired'. We will formalise this in the following section.

As regards the context $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$, the patient's state $\mathcal{S}$ yields initial value-property pairs as facts. With context from Example 3.6, *normalBlood Coagulation*, *normalGastrointestinal Bleeding* $\in \mathcal{S}$ yield facts (in $\mathcal{R}$):

13. *normalBlood Coagulation* $\leftarrow \top$;
14. *normalGastrointestinal Bleeding* $\leftarrow \top$.

Lastly, as in Example 3.6, 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:

- $\mathcal{A} = \{R,\ \mathsf{inapp}(R)\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R}\} \cup \{\mathsf{needs\_repair}(R_i)\ :\ (R_i, R_j, \mathit{Repair}) \in \mathbb{I}\}$ consists of assumptions representing recommendations and those representing the potential non-applicability of recommendations, as well as assumptions representing that recommendations considered repairable need to be repaired;
- $\mathcal{R} = \mathcal{R}_a \cup \mathcal{R}_e \cup \mathcal{R}_v \cup \mathcal{R}_r \cup \mathcal{R}_i \cup \mathcal{R}_p$, where

  - $\mathcal{R}_a = \mathcal{R}_a^+ \cup \mathcal{R}_a^-$ consists of rules representing actions associated with recommendations, where

    * $\mathcal{R}_a^+ = \{A \leftarrow R\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta \geqslant 0\}$,
    * $\mathcal{R}_a^- = \{\mathsf{not}\ A \leftarrow R\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta < 0\}$;

  - $\mathcal{R}_e = \mathcal{R}_e^+ \cup \mathcal{R}_e^-$ consists of rules representing effects on properties brought about by actions, where

    * $\mathcal{R}_e^+ = \{EP \leftarrow A\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta \geqslant 0,\ (P, E, v_I, v_T) \in \mathcal{C}\}$,[16]
    * $\mathcal{R}_e^- = \{\mathsf{not}\ EP \leftarrow \mathsf{not}\ A\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta < 0,\ (P, E, v_I, v_T) \in \mathcal{C}\}$;

  - $\mathcal{R}_v = \mathcal{R}_v^+ \cup \mathcal{R}_v^-$ consists of rules representing the specific values of properties that should (not) be attained given their initial values and the effects brought about by actions, where

    * $\mathcal{R}_v^+ = \{v_T P \leftarrow v_I P, EP\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta \geqslant 0,\ (P, E, v_I, v_T) \in \mathcal{C}\}$,
    * $\mathcal{R}_v^- = \{\mathsf{not}\ v_T P \leftarrow v_I P, \mathsf{not}\ EP\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R},\ \delta < 0,\ (P, E, v_I, v_T) \in \mathcal{C}\}$;

  - $\mathcal{R}_r = \mathcal{R}_{\overline{\mathbb{R}}} \cup \mathcal{R}_{\mathbb{R}}$, where

    * rules in $\mathcal{R}_{\overline{\mathbb{R}}} = \{\overline{R} \leftarrow \mathsf{inapp}(R)\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R}\}$ allow to argue against the default applicability of recommendations if they are inapplicable,
    * rules in $\mathcal{R}_{\mathbb{R}} = \{\overline{\mathsf{inapp}(R)} \leftarrow v_I P\ :\ (R, A, \delta, \mathcal{C}) \in \mathbb{R}, (P, E, v_I, v_T) \in \mathcal{C}\}$ allow to argue against non-applicability (in other words, *for* applicability) of recommendations as long as some property affected can be established with the initial value as per at least one contribution of these recommendations;

  - $\mathcal{R}_i = \mathcal{R}_i^1 \cup \mathcal{R}_i^2$ consists of rules for handling interactions, where

    * rules in $\mathcal{R}_i^1 = \{\overline{R_j} \leftarrow R_k,\ \overline{R_k} \leftarrow R_j\ :\ (R_k, R_j, t) \in \mathbb{I},\ t \in \{\mathit{Contr}, \mathit{Alt}, \mathit{Repet}\}\}$ allow to argue against the default applicability of recommendations given another recommendation in case of contradictions, alternatives or repetitions,[17]
    * rules in $\mathcal{R}_i^2 = \{R_j \leftarrow R_k, \mathsf{needs\_repair}(R_k),\ \overline{\mathsf{needs\_repair}(R_k)} \leftarrow R_j\ :\ (R_k, R_j, \mathit{Repair}) \in \mathbb{I}\}$ allow to argue that following recommendations necessitates following repairing recommendations, as long as repairing is needed, and where arguing against the need for repair is enabled by accepting at least one recommendation that actually does the repair;[18]

  - $\mathcal{R}_p = \{v_I P \leftarrow \top\ :\ v_I P \in \mathcal{S}\}$ consists of facts representing the patient's state $\mathcal{S}$ in terms of properties and their values, where $\mathcal{S} \subseteq \bigcup_{R \in \mathbb{R}} \{v_I P\ :\ (P, E, v_I, v_T) \in \mathcal{C},\ R = (R, A, \delta, \mathcal{C})\}$;

- $\leqslant$ is a preorder over $\mathcal{A}$;
- $\mathcal{G} = \mathcal{G}^+ \cup \mathcal{G}^-$ satisfies

---

[16]Throughout, EP is simply a concatenation of terms; the same applies to value-property pairs.

[17]Note the symmetry of the rules given an interaction, in accordance with the symmetry of elements of $\mathbb{I}$ in the first two components for these three interaction types as discussed in Remark 1.

[18]The asymmetry of these rules is in accordance with Remark 1. Note also that presence of these rules results into non-flatness of the framework.

- $\mathcal{G}^+ \subseteq \bigcup_{R \in \mathbb{R}} \{EP \ : \ (P, E, v_I, v_T) \in \mathcal{C}, \ R = (R, A, \delta, \mathcal{C}), \ \delta \geqslant 0\}$,
- $\mathcal{G}^- \subseteq \bigcup_{R \in \mathbb{R}} \{\text{not } EP \ : \ (P, E, v_I, v_T) \in \mathcal{C}, \ R = (R, A, \delta, \mathcal{C}), \ \delta < 0\}$,

- $\preccurlyeq$ is either empty or 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 $\{\overline{\alpha} \ : \ \alpha \in \mathcal{A}\}$.

Regarding interactions and rules in $\mathcal{R}_i$, on the one hand suppose recommendations R_k and R_j are in a Contradiction, Alternative or Repetition interaction. Then, following R_k is a reason for not following R_j, and vice versa, because, intuitively, they either suggest opposing actions or suggest actions that are interchangeable with respect to their consequences. In other words, following both recommendations would result in either a conflicting, or superfluous (and in many cases undesirable) clinical care situation. On the other hand, if R_j is a recommendation which 'repairs' the consequences of following recommendation R_k, then the latter is a good reason for following the former, assuming R_k 'needs repair', where R_k 'does not need repair' anymore if at least one such repairing recommendation R_j is followed. We discuss some variations of dealing with Repairable interactions in Section 6.2.2.

Finally, we define when a recommendation is applicable.

Definition 4.5. We say that a recommendation $R \in \mathbb{R}$ is *applicable* in the ABA⁺G patient framework $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$ iff $\emptyset \vdash \overline{\text{inapp}(R)}$ is an argument.

Intuitively, a recommendation is applicable if the patient presents with a state in which the recommendation can affect at least one property. Only applicable recommendations are acceptable in ABA⁺G, in the sense that no $<$-preferred extension can contain an inapplicable recommendation, because otherwise it would be $<$-attacked by the empty set, and hence $<$-self-attacking.

Example 4.1. Given recommendations $\mathbb{R}$ from Example 3.1, interactions $\mathbb{I}$ from Example 3.2 and context $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$ from Example 3.6, assumptions $\mathcal{A}$ and rules **1.** to **14.** from Section 4.3.1 specify the ABA⁺G framework $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$ with preferences $\leqslant$, goals $\mathcal{G}$ and priorities $\preccurlyeq$. Due to the patient's state represented as facts, both recommendations R_1 and R_2 are applicable. As they are in Contradiction interaction, we find $\{R_1\} \vdash \overline{R_2}$ and $\{R_2\} \vdash \overline{R_1}$. Due to the preference $R_2 < R_1$, we obtain $\{R_1\} \leadsto_< \{R_2\}$ and $\{R_2\} \not\leadsto_< \{R_1\}$. Since $\{R_1\}$ is clearly closed and $<$-conflict-free, it is plain that it is a unique $<$-preferred extension of $\mathcal{F}_p$. It has, in particular, *Adm. NSAID*, *decreaseBlood Coagulation* and *lowBlood Coagulation* among its conclusions $Cn(\{R_1\})$. Thus, there is also a unique goal extension $\mathcal{G}_{\{R_1\}} = \{decreaseBlood\ Coagulation\}$, which is hence a unique top goal extension. So R_1 is the recommendation suggested by ABA⁺G, in accordance with the patient's preferences.

Now, if the patient had no preferences, i.e. if $\leqslant = RCl(\emptyset)$ instead (see Remark 2 in Section 4.1), both $\{R_1\}$ and $\{R_2\}$ would be (all and only) $<$-preferred extensions. They would induce goal extensions $\mathcal{G}_{\{R_1\}} = \{decreaseBlood\ Coagulation\}$ and $\mathcal{G}_{\{R_2\}} = \{\text{not } increaseGastrointestinal\ Bleeding\}$ such that $\mathcal{G}_{\{R_1\}} \lhd_{\mathbb{G}} \mathcal{G}_{\{R_2\}}$ due to not increasing gastrointestinal bleeding being prioritised over decreasing blood coagulation. Hence, $\mathcal{G}_{\{R_2\}}$ would be a unique top goal extension and R_2 would be the recommendation suggested by ABA⁺G, in accordance with the clinician's prioritised goals.

Having formally defined and illustrated a mapping from TMR recommendations, interactions and the context of the patient to ABA⁺G, we next study some properties of ABA⁺G patient frameworks.

4.3.3. Properties

Modelling recommendations and interactions argumentatively allows to exploit properties of ABA$^+$ to ensure desirable features of our approach. In this section we assume $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^-, \leqslant, \mathcal{G}, \preccurlyeq)$ to be a given ABA$^+$G patient framework. We establish the following properties.

First, $<$-preferred extensions in ABA$^+$G patient frameworks are interaction-free (Definition 3.3) as sets of recommendations (recall that we identify a recommendation with its name, see note after Definition 3.1):

Theorem 4.1 (Interaction-freeness). *For a $<$-preferred extension E of $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^-, \leqslant, \mathcal{G}, \preccurlyeq)$, $E \cap \mathbb{R}$ is an interaction-free set of recommendations.*

Proof. Suppose $E \cap \mathbb{R}$ is not interaction-free. Then, there is $(R_i, R_j, t) \in \mathbb{I}$ with $R_i, R_j \in E$ and $t \in \mathcal{T} \setminus \{Repair\}$. But as $\overline{R_j} \leftarrow R_i \in \mathcal{R}$, we find $E \rightsquigarrow_< E$. This contradicts $<$-conflict-freeness of E. $\square$

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

Corollary 4.2. *For every top goal extension $\mathcal{G}_E$ of $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^-, \leqslant, \mathcal{G}, \preccurlyeq)$ induced by a $<$-preferred extension E, $E \cap \mathbb{R}$ is interaction-free.*

We argue that this property of ABA$^+$G frameworks is desirable, because it ensures that outcomes of reasoning with guidelines and patient information resolve the interactions arising among the applicable recommendations, as intended by the Ariadne principles.

The second property of ABA$^+$G frameworks states that interaction-resolving sets of recommendations (Definition 3.3) are closed and $<$-conflict-free in ABA$^+$G:

Lemma 4.3. *An interaction-resolving set $\mathbb{R}' \subseteq \mathbb{R}$ of recommendations is closed and $<$-conflict-free in $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, {}^-, \leqslant, \mathcal{G}, \preccurlyeq)$.*

Proof. Consider $\alpha \in Cn(\mathbb{R}') \cap \mathcal{A}$. By construction of $\mathcal{F}_p$, the only way α could be in $Cn(\mathbb{R}')$ but not in $\mathbb{R}'$ is when $\alpha = R_j \in \mathbb{R}$ and there is $R_j \leftarrow R_i$, needs_repair$(R_i) \in \mathcal{R}$ such that R_i, needs_repair$(R_i) \in \mathbb{R}'$. But since needs_repair$(R_i) \notin Cn(\mathbb{R}')$, this cannot happen, and so it must be that $Cn(\mathbb{R}') \cap \mathcal{A} \subseteq \mathbb{R}'$. Since trivially $\mathbb{R}' \subseteq Cn(\mathbb{R}') \cap \mathcal{A}$, we find that $\mathbb{R}' = Cn(\mathbb{R}') \cap \mathcal{A}$, i.e. $\mathbb{R}'$ is closed.

Since $\mathbb{R}'$ is interaction-resolving, it is by definition interaction-free. As $\mathbb{R}'$ is closed, the only way it can be non-$<$-conflict-free is when there is $\overline{R_j} \leftarrow R_i \in \mathcal{R}$ such that $R_i, R_j \in \mathbb{R}'$. But the existence of such rule entails $(R_i, R_j, t) \in \mathbb{I}$ with $t \in \{Contr, Alt, Repet\}$. This, however, contradicts $\mathbb{R}'$ being interaction-free. Hence, $\mathbb{R}'$ must be $<$-conflict-free. $\square$

Since closure and $<$-conflict-freeness are fundamental requirements for any semantics in ABA$^+$, Lemma 4.3 ensures that interaction-resolving recommendations meet the fundamental requirements for acceptance in ABA$^+$G.

Another property states that if the patient expresses preferences over all recommendations, then the most preferred interaction-resolving applicable recommendations (see Definition 4.5) will be followed:

Theorem 4.4 (**Preferences Theorem**). *Let $\leqslant$ be a total order over $\mathbb{R}$ and suppose that the set $\mathbb{R}' = \{R' \in \mathbb{R} : R'$ is applicable and there is no applicable $R \in \mathbb{R}$ with $R' < R\}$ of the most preferred applicable recommendations is interaction-resolving. Then, for every $<$-preferred extension E of $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A},^-, \leqslant, \mathcal{G}, \preccurlyeq)$ it holds that $\mathbb{R}' \subseteq E$.*

Proof. Let $E \subseteq \mathcal{A}$ be a $<$-preferred extension of $\mathcal{F}_p$. By Lemma 4.3, $\mathbb{R}'$ is closed and $<$-conflict-free. This, with $\mathbb{R}'$ consisting of $\leqslant$-maximal applicable recommendations, entails that $\mathbb{R}'$ is $<$-defended by E, as follows.

- Concerning reverse attacks (see Section 4.1), as recommendations in $\mathbb{R}'$ are $\leqslant$-maximal applicable, by construction of $\mathcal{F}_p$, $\mathbb{R}'$ can be $<$-attacked via reverse attack only if $\{R'\} \vdash \overline{R}$ for some $R' \in \mathbb{R}'$ and $R \in \mathbb{R}$ such that $R' < R$. As $\mathbb{R}'$ is $<$-conflict-free, we have that $R \notin \mathbb{R}'$. But then $R' \in \mathbb{R}'$ and $R' < R$ entail that R is not applicable. Hence, there is no argument $\emptyset \vdash \overline{\mathsf{inapp}(R)}$, and so by construction of $\mathcal{F}_p$, $\{\mathsf{inapp}(R)\}$ is $<$-unattacked. From here, utilising the construction of $\mathcal{F}_p$, we show that $\mathsf{inapp}(R) \in E$, whence $E <$-defends $\{R'\}$. So suppose for a contradiction that $\mathsf{inapp}(R) \notin E$. We claim that $E \cup \{\mathsf{inapp}(R)\}$ is $<$-admissible. Indeed, by construction of $\mathcal{R}$, $E \cup \{\mathsf{inapp}(R)\}$ is clearly closed, as E is closed. It is plain to see that the only way it would not be $<$-conflict-free necessitates $R \in E$. But in that case, E could not $<$-defend against $\{\mathsf{inapp}(R)\} \rightsquigarrow_< \{R\}$ due to R being inapplicable. Hence, by contradiction, $E \cup \{\mathsf{inapp}(R)\}$ is $<$-conflict-free, and hence $<$-admissible. But this contradicts E being $<$-preferred. Thus, by contradiction, $\mathsf{inapp}(R) \in E$. It then follows that $E <$-defends $\{R'\}$ against reverse attacks. Consequently, $E <$-defends $\mathbb{R}'$ against reverse attacks.
- Concerning normal attacks, suppose for a contradiction that for some $A \subseteq \mathcal{A}$ it holds that $A \vdash \overline{R'}$ and $\forall \alpha \in A \ \alpha \not< R'$. By construction of $\mathcal{F}_p$, the contrary of an assumption can only be deduced from another assumption, so that A is a singleton taking one of the two forms below.
    - First suppose $A = \{\mathsf{inapp}(R')\}$. Since R' is applicable, we have $\emptyset \vdash \overline{\mathsf{inapp}(R')}$. Since $\emptyset \subseteq E$, $E <$-defends against A.
    - Now suppose $A = \{R\}$ for some $R \in \mathbb{R}$. Then $R \not< R'$, and since $\leqslant$ is total over $\mathbb{R}$, it holds that $R' \leqslant R$. But this means that R is either not applicable or $R \in \mathbb{R}'$: indeed, if R is applicable and $R \notin \mathbb{R}'$, then there is some applicable $R'' \in \mathbb{R}$ such that $R < R''$, whence by transitivity of $\leqslant$ we have $R' < R''$ too, which contradicts R' being $\leqslant$-maximal applicable. But note also that if $R \in \mathbb{R}'$, then $\mathbb{R}'$ is not $<$-conflict-free, and hence not interaction-resolving, according to Lemma 4.3, which is a contradiction. Hence, R is inapplicable, and so there is no argument $\emptyset \vdash \overline{\mathsf{inapp}(R)}$, which again means that $\{\mathsf{inapp}(R)\}$ is $<$-unattacked and $\mathsf{inapp}(R) \in E$. In turn, $E <$-defends $\{R'\}$ against A.

  Thus, $E <$-defends $\{R'\}$ against normal attacks, and so it $<$-defends $\mathbb{R}'$ against normal attacks.

So $E <$-defends $\mathbb{R}'$. Suppose then for a contradiction that $\mathbb{R}' \nsubseteq E$. We show that $E \cup \mathbb{R}'$ is $<$-admissible.

- Since both E and $\mathbb{R}'$ are closed, the only way their union would not be closed is, by construction of $\mathcal{F}_p$, if there were $R_j \leftarrow R_i, \mathsf{needs\_repair}(R_i) \in \mathcal{R}$ such that $R_i \in \mathbb{R}'$, $\mathsf{needs\_repair}(R_i) \in E$ and $R_j \notin E \cup \mathbb{R}'$. As $\mathbb{R}'$ is interaction-resolving, it holds that there would be $R_k \in \mathbb{R}'$ with $(R_i, R_k, Repair) \in \mathbb{I}$. We would then have $\{R_k\} \vdash \overline{\mathsf{needs\_repair}(R_i)}$ so that $\{R_k\} \rightsquigarrow_< \{\mathsf{needs\_repair}(R_i)\}$. But then, E being $<$-preferred and $\mathsf{needs\_repair}(R_i) \in E$ would imply that $E <$-attacks $\{R_k\}$. But since $R_k \in \mathbb{R}'$ and we established that $E <$-defends $\mathbb{R}'$, this would lead to a contradiction to E being $<$-conflict-free. Thus, by contradiction, $E \cup \mathbb{R}'$ must be closed.

- If $E \cup \mathbb{R}'$ is not $<$-conflict-free, then, as it is closed, it must be that either $E \rightsquigarrow_< \mathbb{R}'$ or $\mathbb{R}' \rightsquigarrow_< E$. Either case necessitates $E \rightsquigarrow_< E$, because E $<$-defends $\mathbb{R}'$. This leads to a contradiction to E being $<$-conflict-free. Thus, by contradiction, $E \cup \mathbb{R}'$ is $<$-conflict-free.
- Since E is $<$-preferred, $<$-defends the closed $\mathbb{R}'$ and $E \cup \mathbb{R}'$ is closed, E $<$-defends $E \cup \mathbb{R}'$ too. Hence, $E \cup \mathbb{R}'$ is $<$-admissible.

We thus obtain a contradiction to E being a $<$-preferred extension. Therefore, by contradiction it holds that $\mathbb{R}' \subseteq E$, as required. □

As priorities over goals are used to select among goal extensions induced by $<$-preferred extensions, top goal extensions (under the same conditions) are obtained by following the most preferred interaction-resolving recommendations:

Corollary 4.5. *Let $\leqslant$ be total and $\mathbb{R}' = \{R' \in \mathbb{R} : R'$ is $\leqslant$-maximal applicable$\}$ interaction-resolving. Then, for every top goal extension $\mathcal{G}_E$ of $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, \bar{}, \leqslant, \mathcal{G}, \preccurlyeq)$ induced by a $<$-preferred extension E, it holds that $\mathbb{R}' \subseteq E$.*

Proof. Follows from Theorem 4.4, since top goal extensions are selected among the goal extensions induced by the $<$-preferred extensions of $\mathcal{F}_p$. □

We argue that this property of ABA⁺G frameworks is desirable, because it ensures that the patient's most preferred recommendations, if applicable, are returned as part of the outcomes of reasoning with guideline recommendations, as intended by the Ariadne principles.

In general, Theorems 4.1 and 4.4, together with Corollaries 4.2 and 4.5, pave the way for ABA⁺G to meet the three Ariadne principles of *interaction assessment*, *prioritisation and patient preferences* and *individualised management* when applied to patient-centric reasoning with conflicting medical recommendations.

5. Illustration

We exemplify the use of ABA⁺G with a case study from [108], focusing on interactions among Breast Cancer (BC), Osteoarthritis (OA), Hypertension (HT) and Congestive Heart Failure (CHT) guidelines. Graphical representation of the relevant guideline recommendations as well as interactions thereof are depicted in Fig. 2, as taken from [108, Fig. 5, p. 87]. (The underlying details can be found in [108, p. 90, Table 9, p. 91, Table 10].) Our results will accord with the informal discussion on the case study in [108].

5.1. Adaptation of case study

Dictated by the design choices of ABA⁺G, we make the following adaptations regarding recommendations and interactions used in this case study.

1. First, we assume that the deontic strengths of all the recommendations are discretised as *should* (*should not*). This means that instead of *must* and *must not*, recommendations R_1, R_5 and R_4 have deontic strengths *should*, *should* and *should not*, respectively. This assumption does not influence the existence of interactions, but would only influence their modal strength, which we do not use in this paper (see Section 3.1.1). Hence, in our setting, this assumption is made without the loss of generality.
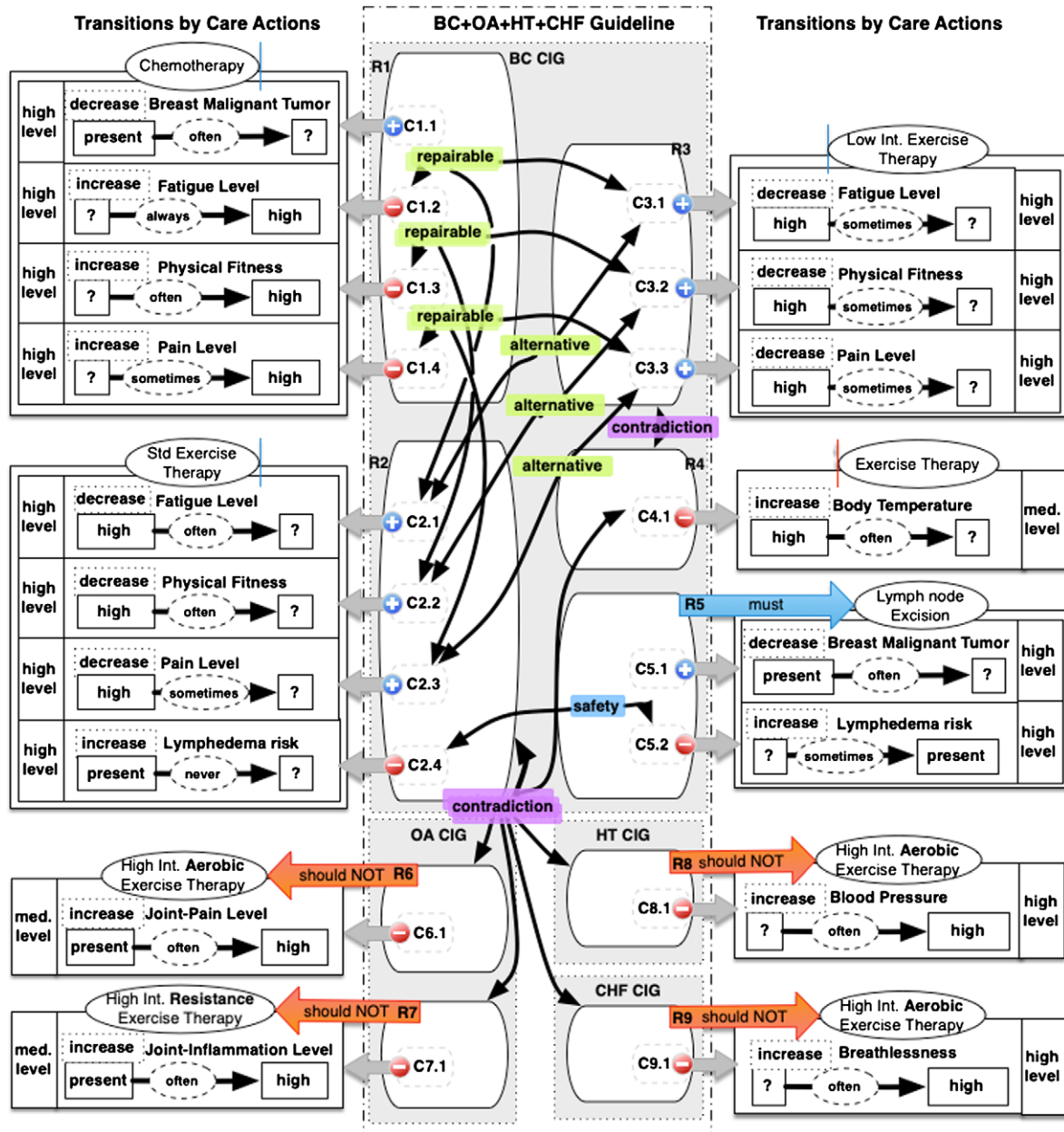
Fig. 2. TMR recommendations and interactions in case study for a merged breast cancer, osteoarthritis, hypertension and congestive heart failure guideline [108, p. 87, Fig. 5]. The black arrows indicate which recommendations (or more precisely, contributions) are in an interaction of the specified type. (Figure kindly provided by the authors of [108].)

2. Second, as already mentioned in footnote 9, we instantiate all the indeterminate values with values such as *residual*, *normal*, *very high*, *elevated*. This does not affect the reasoning outcomes in terms of <-preferred and top goal extensions in ABA⁺G, and so does not result into loss of generality.

3. Further, as in [108], a hierarchy of actions associated to recommendations is assumed (cf. footnote 10). In particular, the actions concerning various exercises are related in the following

way. (a) *Std. Exercise* consists of *High Int. Aerobic Exercise* and *High Int. Resistance Exercise*.[19]
(b) *Exercise* pertains to *any* available exercise therapy or any combination thereof. That is, it
can be e.g. *High Int. Aerobic Exercise* on its own, or *Std. Exercise*, or *Low Int. Exercise*. (c)
Low Int. Exercise is not further specified. This hierarchy is specified in TMRweb and is used to
detect interactions among recommendations. Since ABA$^+$G takes interactions afforded by TMR
as input, the action hierarchy is not needed for reasoning purposes. Thus, this assumption in our
setting is also made without the loss of generality.

4. Finally, note that contribution $C_{2.4}$ of recommendation R_2 has causation probability (see Section 3.1.1) *never*, meaning that R_2 never leads to an increase to *Lymphoedema Risk*. Since this contribution is relevant only to the so-called "safety" interactions which we do not address in this paper, this simplifying assumption does not result in the loss of generality in our setting.

5.2. Recommendations and interactions for ABA$^+$G

We now spell out the (adapted) recommendations and interactions thereof appearing in this case study.
Let $\mathbb{R} = \{R_i \ : \ i \in \{1, \ldots, 9\}\}$ with the following recommendations.

- (R_1, *Chemotherapy*, *should*, {

    (*Breast Tumour*, *decrease*, *present*, *residual*),
    (*Fatigue*, *increase*, *normal*, *high*),
    (*Fitness*, *increase*, *normal*, *high*),
    (*Pain*, *increase*, *normal*, *high*)

    })
- (R_2, *Std. Exercise*, *should*, {

    (*Fatigue*, *decrease*, *high*, *normal*),
    (*Fitness*, *decrease*, *high*, *normal*),
    (*Pain*, *decrease*, *high*, *normal*)

    })
- (R_3, *Low Int. Exercise*, *should*, {

    (*Fatigue*, *decrease*, *high*, *normal*),
    (*Fitness*, *decrease*, *high*, *normal*),
    (*Pain*, *decrease*, *high*, *normal*)

    })
- (R_4, *Exercise*, *should not*, {(*Body Temperature*, *increase*, *high*, *very high*)})
- (R_5, *Lymph node Excision*, *should*, {

    (*Breast Tumour*, *decrease*, *present*, *residual*),
    (*Lymphoedema Risk*, *increase*, *present*, *present*)

    })
- (R_6, *High Int. Aerobic Exercise*, *should not*, {(*Joint-Pain*, *increase*, *present*, *high*)})
- (R_7, *High Int. Resistance Exercise*, *should not*, {(*Joint-Inflammation*, *increase*, *present*, *high*)})

---

[19]We omit the word 'Therapy' for concision.

- (R_8, *High Int. Aerobic Exercise*, *should not*, {(*Blood Pressure*, *increase*, *elevated*, *high*)})
- (R_9, *High Int. Aerobic Exercise*, *should not*, {(*Breathlessness*, *increase*, *present*, *high*)})

We then have the following interactions.

$$\mathbb{I} = \big\{ (R_2, R_4, Contr), (R_2, R_6, Contr), (R_2, R_7, Contr), (R_2, R_8, Contr), (R_2, R_9, Contr),$$
$$(R_3, R_4, Contr), (R_2, R_3, Alt), (R_1, R_2, Repair), (R_1, R_3, Repair) \big\}$$

Without any information about patients, $(\mathbb{R}, \mathbb{I})$ will result into the following assumptions and rules in ABA⁺G.

- $A = \big\{ R_i, \ \mathsf{inapp}(R_i) \ : \ i \in \{1, \dots, 9\} \big\} \cup \big\{ \mathsf{needs\_repair}(R_1) \big\}$,

- $\mathcal{R}_a = \{$*Chemotherapy* $\leftarrow R_1$, *Std. Exercise* $\leftarrow R_2$, *Low Int. Exercise* $\leftarrow R_3$,

     not *Exercise* $\leftarrow R_4$, *Lymph node Excision* $\leftarrow R_5$,

     not *High Int. Aerobic Exercise* $\leftarrow R_6$, not *High Int. Resistance Exercise* $\leftarrow R_7$,

     not *High Int. Aerobic Exercise* $\leftarrow R_8$, not *High Int. Aerobic Exercise* $\leftarrow R_9\}$,

- $\mathcal{R}_e = \{$*decreaseBreast Tumour* $\leftarrow$ *Chemotherapy*, *increaseFatigue* $\leftarrow$ *Chemotherapy*,

     increaseFitness $\leftarrow$ *Chemotherapy*, *increasePain* $\leftarrow$ *Chemotherapy*,

     decreaseFatigue $\leftarrow$ *Std. Exercise*, *decreaseFitness* $\leftarrow$ *Std. Exercise*,

     decreasePain $\leftarrow$ *Std. Exercise*, *decreaseFatigue* $\leftarrow$ *Low Int. Exercise*,

     decreaseFitness $\leftarrow$ *Low Int. Exercise*, *decreasePain* $\leftarrow$ *Low Int. Exercise*,

     not *increaseBody Temperature* $\leftarrow$ not *Exercise*,

     decreaseBreast Tumour $\leftarrow$ *Lymph node Excision*,

     increaseLymphoedema Risk $\leftarrow$ *Lymph node Excision*,

     not *increaseJoint-Pain* $\leftarrow$ not *High Int. Aerobic Exercise*,

     not *increaseJoint-Inflammation* $\leftarrow$ not *High Int. Resistance Exercise*,

     not *increaseBlood Pressure* $\leftarrow$ not *High Int. Aerobic Exercise*,

     not *increaseBreathlessness* $\leftarrow$ not *High Int. Aerobic Exercise*$\}$,

- $\mathcal{R}_v = \{$*residualBreast Tumour* $\leftarrow$ *presentBreast Tumour*, *decreaseBreast Tumour*,

     highFatigue $\leftarrow$ *normalFatigue*, *increaseFatigue*, *highPain* $\leftarrow$ *normalPain*, *increasePain*,

     highFitness $\leftarrow$ *normalFitness*, *increaseFitness*,

     normalFatigue $\leftarrow$ *highFatigue*, *decreaseFatigue*,

     normalPain $\leftarrow$ *highPain*, *decreasePain*,

     normalFitness $\leftarrow$ *highFitness*, *decreaseFitness*,

     not *very highBody Temperature* $\leftarrow$ *highBody Temperature*,

not *increaseBody Temperature*,

presentLymphoedema Risk ← presentLymphoedema Risk, increaseLymphoedema Risk,

not *highJoint-Pain ← presentJoint-Pain*, not *High Int. Aerobic Exercise*,

not *highJoint-Inflammation ← presentJoint-Inflammation*,

not *High Int. Resistance Exercise*,

not *highBlood Pressure ← elevatedBlood Pressure*, not *High Int. Aerobic Exercise*,

not *highBreathlessness ← presentBreathlessness*, not *High Int. Aerobic Exercise*},

- $\mathcal{R}_r = \left\{ \overline{R_i} \leftarrow \mathsf{inapp}(R_i) \; : \; i \in \{1, \ldots, 9\} \right\}$

  $\cup \; \left\{ \overline{\mathsf{inapp}(R_1)} \leftarrow presentBreast\ Tumour, \right.$

  $\overline{\mathsf{inapp}(R_1)} \leftarrow normalFatigue,$

  $\overline{\mathsf{inapp}(R_1)} \leftarrow normalFitness, \; \overline{\mathsf{inapp}(R_1)} \leftarrow normalPain,$

  $\overline{\mathsf{inapp}(R_2)} \leftarrow highFatigue, \; \overline{\mathsf{inapp}(R_2)} \leftarrow highFitness, \; \overline{\mathsf{inapp}(R_2)} \leftarrow highPain,$

  $\overline{\mathsf{inapp}(R_3)} \leftarrow highFatigue, \; \overline{\mathsf{inapp}(R_3)} \leftarrow highFitness, \; \overline{\mathsf{inapp}(R_3)} \leftarrow highPain,$

  $\overline{\mathsf{inapp}(R_4)} \leftarrow highBody\ Temperature,$

  $\overline{\mathsf{inapp}(R_5)} \leftarrow presentBreast\ Tumour, \; \overline{\mathsf{inapp}(R_5)} \leftarrow presentLymphoedema\ Risk,$

  $\overline{\mathsf{inapp}(R_6)} \leftarrow presentJoint-Pain, \; \overline{\mathsf{inapp}(R_7)} \leftarrow presentJoint-Inflammation,$

  $\left. \overline{\mathsf{inapp}(R_8)} \leftarrow elevatedBlood\ Pressure, \; \overline{\mathsf{inapp}(R_9)} \leftarrow presentBreathlessness \right\},$

- $\mathcal{R}_i = \{ \overline{R_2} \leftarrow R_4, \; \overline{R_4} \leftarrow R_2, \; \overline{R_2} \leftarrow R_6, \; \overline{R_6} \leftarrow R_2, \; \overline{R_2} \leftarrow R_7, \; \overline{R_7} \leftarrow R_2, \; \overline{R_2} \leftarrow R_8,$

  $\overline{R_8} \leftarrow R_2, \; \overline{R_2} \leftarrow R_9, \; \overline{R_9} \leftarrow R_2, \; \overline{R_3} \leftarrow R_4, \; \overline{R_4} \leftarrow R_3, \; \overline{R_2} \leftarrow R_3, \; \overline{R_3} \leftarrow R_2 \}$

  $\cup \; \left\{ R_2 \leftarrow R_1, \mathsf{needs\_repair}(R_1), \; R_3 \leftarrow R_1, \mathsf{needs\_repair}(R_1), \right.$

  $\left. \overline{\mathsf{needs\_repair}(R_1)} \leftarrow R_2, \; \overline{\mathsf{needs\_repair}(R_1)} \leftarrow R_3 \right\}.$

5.3. Patients and their contexts

We will consider four patients in the following contexts.

1. We assume all patients to have Breast Cancer present and exhibit fatigue. To illustrate a basic scenario of reasoning with guideline recommendations and interactions, we assume that the first patient has no preferences and also that no explicit goals are set for them:
   - $\mathcal{S}_1 = \{presentBreast\ Tumour, \; highFatigue\}$,
   - $\mathcal{G}_1 = \emptyset$,
   - $\leqslant_1 = RCl(\emptyset) = \{(a, a) \; : \; a \in \mathcal{A}\}$,[20]

---

[20]$RCl(\emptyset)$ is the reflexive closure of $\emptyset$; see Remark 2.

- $\preccurlyeq_1 = \emptyset$.

2. We assume the second patient to have comorbidity Osteoarthritis and so complain of joint-pain and joint-inflammation. They thus express the preference to not exercise intensely:

   - $\mathcal{S}_2 = \mathcal{S}_1 \cup \{presentJoint\text{-}Pain, \ presentJoint\text{-}Inflammation\}$,
   - $\mathcal{G}_2 = \mathcal{G}_1 = \emptyset$,
   - $\leqslant_2$ given by (the reflexive and transitive closure of $<_2$ with) $R_2 <_2 R_6$, $R_2 <_2 R_7$,
   - $\preccurlyeq_2 = \preccurlyeq_1 = \emptyset$.

3. Patient 3, in addition to Breast Cancer, has Hypertension and thus suffers from elevated blood pressure. Accordingly, the clinician sets the goals to first and foremost alleviate the breast tumour, followed by not sending the blood pressure even higher, and, if possible, relieving the anticipated fatigue, fitness and pain issues:

   - $\mathcal{S}_3 = \mathcal{S}_1 \cup \{elevatedBlood\ Pressure\}$,
   - $\mathcal{G}_3 = \{decreaseBreast\ Tumour, \ \mathsf{not}\ increaseBlood\ Pressure,$
        $decreaseFatigue, \ decreaseFitness, \ decreasePain\}$,
   - $\leqslant_3 = \leqslant_1$,
   - $\preccurlyeq_3$ given by

     – $decreaseFatigue \prec_3 \mathsf{not}\ increaseBlood\ Pressure \prec_3 decreaseBreast\ Tumour$,
     – $decreaseFitness \prec_3 \mathsf{not}\ increaseBlood\ Pressure \prec_3 decreaseBreast\ Tumour$,
     – $decreasePain \prec_3 \mathsf{not}\ increaseBlood\ Pressure \prec_3 decreaseBreast\ Tumour$,
     – $decreaseFatigue \preccurlyeq_3 decreaseFitness \preccurlyeq_3 decreasePain \preccurlyeq_3 decreaseFatigue$,

   visualised below.

$$decreaseBreast\ Tumour$$
$$|$$
$$\mathsf{not}\ increaseBlood\ Pressure$$

$$decreaseFatigue \qquad decreaseFitness \qquad decreasePain$$
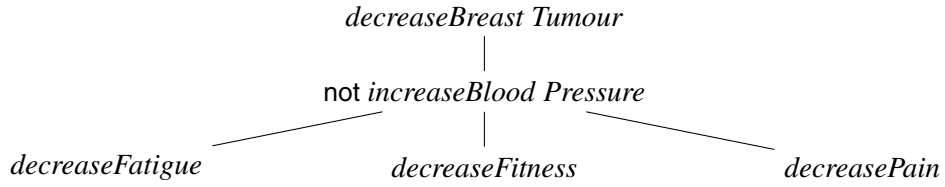
4. The fourth patient is in the same situation as the third one, but also has a preference for low intensity exercise therapy (as per recommendation R_3) over standard exercise (as per R_2):

   - $\mathcal{S}_4 = \mathcal{S}_3$,
   - $\mathcal{G}_4 = \mathcal{G}_3$,
   - $\leqslant_4$ given by $R_2 <_4 R_3$,
   - $\preccurlyeq_4 = \preccurlyeq_3$.

We are now ready to construct four ABA⁺G patient frameworks $\mathcal{F}_k = (\mathcal{L}_k, \mathcal{R}_k, \mathcal{A},^-, \leqslant_k, \mathcal{G}_k, \preccurlyeq_k)$, for $k \in \{1, 2, 3, 4\}$, with $\leqslant_k, \mathcal{G}_k$ and $\preccurlyeq_k$ as in every patient's context $(\mathcal{S}_k, \mathcal{G}_k, \leqslant_k, \preccurlyeq_k)$ above, and where $\mathcal{R}_k = \mathcal{R}_a \cup \mathcal{R}_e \cup \mathcal{R}_r \cup \mathcal{R}_i \cup \mathcal{R}_{p_k}$ with

- $\mathcal{R}_{p_1} = \{presentBreast\ Tumour \leftarrow \top, \ highFatigue \leftarrow \top\}$,
- $\mathcal{R}_{p_2} = \mathcal{R}_{p_1} \cup \{presentJoint\text{-}Pain \leftarrow \top, \ presentJoint\text{-}Inflammation \leftarrow \top\}$,
- $\mathcal{R}_{p_3} = \mathcal{R}_{p_1} \cup \{elevatedBlood\ Pressure \leftarrow \top\}$,
- $\mathcal{R}_{p_4} = \mathcal{R}_{p_3}$.

5.4. Patient-centric reasoning in ABA$^+$G

We now execute patient-centric reasoning with interacting clinical guideline recommendations in ABA$^+$G.

5.4.1. Patient 1

Consider first $\mathcal{F}_1$. Note immediately that $\nexists A \vdash \overline{\mathsf{inapp}(R_i)}$ with $A \subseteq \mathcal{A}$ for $i \in \{4, 6, 7, 8, 9\}$. Thus, recommendations R_4, R_6, R_7, R_8, R_9 are inapplicable, and hence will not belong to any $<$-preferred extension of $\mathcal{F}_1$. The remaining recommendations are all applicable, because presence of the fact *presentBreast Tumour* $\leftarrow \top$ and rules $\overline{\mathsf{inapp}(R_1)} \leftarrow$ *presentBreast Tumour*, $\overline{\mathsf{inapp}(R_5)} \leftarrow$ *presentBreast Tumour* in $\mathcal{R}_1$ give arguments $\emptyset \vdash \overline{\mathsf{inapp}(R_1)}$ and $\emptyset \vdash \overline{\mathsf{inapp}(R_5)}$; similarly, *highFatigue* $\leftarrow \top \in \mathcal{R}_1$ leads to $\emptyset \vdash \overline{\mathsf{inapp}(R_2)}$ and $\emptyset \vdash \overline{\mathsf{inapp}(R_3)}$. (Here we focus on recommendation applicability within ABA$^+$G; in Section 6.2.1 we discuss how this could be accounted for differently, outside ABA$^+$G and within a deployed decision support system.)

So concerning the remaining applicable recommendations, note first that R_5 is not involved in any interaction, so that $\{R_5\}$ is not $<$-attacked by any set not containing $\mathsf{inapp}(R_5)$ (where $\emptyset \leadsto_< \{\mathsf{inapp}(R_5)\}$). Thus, R_5 will belong to every $<$-preferred extension. On the other hand, we have $\{R_2\}$ and $\{R_3\}$ $<$-attacking each other, because R_2 and R_3 are in Alternative interaction and there are no preferences between the two recommendations. In addition, both can 'repair' R_1. Since $\{R_1\}$ is not $<$-attacked by any set not containing $\mathsf{inapp}(R_1)$, R_1 could be accepted, depending on the acceptance of $\mathsf{needs\_repair}(R_1)$.

a) Note first that $\{R_1, \mathsf{needs\_repair}(R_1)\}$ is not closed. This means that accepting both R_1 and $\mathsf{needs\_repair}(R_1)$ necessitates accepting everything in the closure $Cl(\{R_1, \mathsf{needs\_repair}(R_1)\})$, particularly R_2 and R_3. However, both $\{R_2\}$ and $\{R_3\}$ $<$-attack $\{\mathsf{needs\_repair}(R_1)\}$. So $Cl(\{R_1, \mathsf{needs\_repair}(R_1)\})$ is actually self-$<$-attacking, and hence not acceptable.

b) $\mathsf{needs\_repair}(R_1)$ on its own cannot be accepted due to $<$-attacks $\{R_2\} \leadsto_< \{\mathsf{needs\_repair}(R_1)\}$ and $\{R_3\} \leadsto_< \{\mathsf{needs\_repair}(R_1)\}$. But as only $\{R_3\}$ and $\{R_2\}$ can $<$-defends against $\{R_2\}$ and $\{R_3\}$, respectively, $\mathsf{needs\_repair}(R_1)$ could be accepted only alongside either R_2 or R_3, which is again impossible due to the above $<$-attacks.

In the end, R_1 and R_5 should be accepted on their own. In addition, as $<$-preferred extensions need to be $\subseteq$-maximally $<$-admissible, exactly one of R_2 and R_3 should be accepted too. Therefore, $\{R_1, R_2, R_5\}$ and $\{R_1, R_3, R_5\}$ are all and only $<$-preferred extensions of $\mathcal{F}_1$, with conclusions:

- $Cn(\{R_1, R_2, R_5\}) = \{$*presentBreast Tumour*, *highFatigue*, R_1, R_2, R_5,   (i)

  Chemotherapy, *Std. Exercise*, *Lymph node Excision*,   (ii)

  decreaseBreast Tumour, *increaseFatigue*, *increaseFitness*, *increasePain*,   (iii)

  decreaseFatigue, *decreaseFitness*, *decreasePain*, *increaseLymphoedema Risk*,   (iv)

  residualBreast Tumour, *normalFatigue*,   (v)

  $\overline{R_3}, \overline{R_4}, \overline{R_6}, \overline{R_7}, \overline{R_8}, \overline{R_9}, \overline{\mathsf{inapp}(R_1)}, \overline{\mathsf{inapp}(R_2)}, \overline{\mathsf{inapp}(R_3)}, \overline{\mathsf{inapp}(R_5)},$ (vi)

  $\mathsf{inapp}(R_4), \mathsf{inapp}(R_6), \mathsf{inapp}(R_7), \mathsf{inapp}(R_8), \mathsf{inapp}(R_9), \overline{\mathsf{needs\_repair}(R_1)}\},$ (vii)

- $Cn(\{R_1, R_3, R_5\}) = \{presentBreast\ Tumour,\ highFatigue,\ R_1,\ R_3,\ R_5,$ (i)

     Chemotherapy, Low Int. Exercise, Lymph node Excision,   (ii)

     decreaseBreast Tumour, increaseFatigue, increaseFitness, increasePain,   (iii)

     decreaseFatigue, decreaseFitness, decreasePain, increaseLymphoedema Risk,   (iv)

     residualBreast Tumour, normalFatigue,   (v)

     $\overline{R_2},\ \overline{R_4},\ \overline{\mathsf{inapp}(R_1)},\ \overline{\mathsf{inapp}(R_2)},\ \overline{\mathsf{inapp}(R_3)},\ \overline{\mathsf{inapp}(R_5)},$ (vi)

     $\mathsf{inapp}(R_4),\ \mathsf{inapp}(R_6),\ \mathsf{inapp}(R_7),\ \mathsf{inapp}(R_8),\ \mathsf{inapp}(R_9),\ \overline{\mathsf{needs\_repair}(R_1)}\}.$ (vii)

In accordance with Theorem 4.1, both extensions are interaction-free sets of recommendations. The conclusions of either extension indicate (i) the state the patient is already in and the suggested recommendations, (ii) the actions implied by the suggested recommendations, (iii–v) the foreseen consequences of those actions, and (vi–vii) other information pertaining to other available recommendations. In particular, ABA$^+$G suggests that Patient 1 should undergo chemotherapy as well as lymph node excision, accompanied by either standard or low intensity exercise therapies, in order to alleviate Breast Cancer and ease fatigue.

5.4.2. Patient 2

We now consider $\mathcal{F}_2$. Here, R_4 is inapplicable as in $\mathcal{F}_1$, but R_6 and R_7 are applicable, due to patient's Osteoarthritis. We also have $\{R_6\} \vdash \overline{R_2}$ and $\{R_7\} \vdash \overline{R_2}$, and since both R_6 and R_7 are preferred over R_2, we have $<$-attacks ($\{R_6\} \rightsquigarrow_< \{R_2\}$ and $\{R_7\} \rightsquigarrow_< \{R_2\}$) against which $\{R_2\}$ does not $<$-defend. Since $\{R_6\}$ and $\{R_7\}$ are otherwise not $<$-attacked (by any sets not containing either $\mathsf{inapp}(R_6)$ or $\mathsf{inapp}(R_7)$, respectively), this entails that both recommendations will be accepted. Regarding the rest, similarly to $\mathcal{F}_1$, R_1, R_3 and R_5 will be accepted too. So $\mathcal{F}_2$ has a unique $<$-preferred extension $\{R_1, R_3, R_5, R_6, R_7\}$, with conclusions:

- $Cn(\{R_1, R_3, R_5, R_6, R_7\}) = \{presentBreast\ Tumour,\ highFatigue,$

     presentJoint-Pain, presentJoint-Inflammation, $R_1,\ R_3,\ R_5,\ R_6,\ R_7,$

     Chemotherapy, Lymph node Excision, Low Int. Exercise,

     not *High Int. Aerobic Exercise,* not *High Int. Resistance Exercise,*

     decreaseBreast Tumour, increaseFatigue, increaseFitness, increasePain,

     increaseLymphoedema Risk, decreaseFatigue, decreaseFitness, decreasePain,

     not *increaseJoint-Pain,* not *increaseJoint-Inflammation,*

     residualBreast Tumour, normalFatigue, not *highJoint-Pain,* not *highJoint-Inflammation,*

     $\overline{R_2},\ \overline{R_4},\ \overline{\mathsf{inapp}(R_1)},\ \overline{\mathsf{inapp}(R_2)},\ \overline{\mathsf{inapp}(R_3)},\ \overline{\mathsf{inapp}(R_5)},\ \overline{\mathsf{inapp}(R_6)},\ \overline{\mathsf{inapp}(R_7)},$

     $\mathsf{inapp}(R_4),\ \mathsf{inapp}(R_8),\ \mathsf{inapp}(R_9),\ \overline{\mathsf{needs\_repair}(R_1)}\}.$

Note that $\{R_1, R_3, R_5, R_6, R_7\}$ is actually an interaction-resolving set of the most preferred applicable recommendations, and so, in accordance with Theorem 4.4, it is (contained in) the unique $<$-preferred

extension of $\mathcal{F}_2$. The suggested recommendations for Patient 2 are hence to undergo chemotherapy, lymph node excision and to engage in low intensity exercise, as well as, explicitly, in accordance with the patient's preferences, not to engage in high intensity (aerobic and resistance) exercise so as not to worsen joint-pain and inflammation.

5.4.3. Patient 3

In $\mathcal{F}_3$, the only applicable recommendation in addition to those applicable in $\mathcal{F}_1$ is R_8, due to patient's Hypertension resulting into blood pressure problems: in contrast to $\mathcal{F}_1$, we find $\emptyset \vdash \overline{\mathsf{inapp}(R_8)}$ due to the presence of the rule $\overline{\mathsf{inapp}(R_8)} \leftarrow elevatedBlood\,Pressure$ and the patient having blood pressure issues as encoded by the rule $elevatedBlood\,Pressure \leftarrow \top$.

In addition to the relevant $<$-attacks due to interactions as in $\mathcal{F}_1$, we have $\{R_8\}$ and $\{R_2\}$ $<$-attacking each other. Since $\{R_8\}$ is otherwise not $<$-attacked (by any set not containing $\mathsf{inapp}(R_8)$), R_8 could be accepted, alongside R_1, R_3 and R_5. On the other hand, R_1, R_2 and R_5 are still collectively acceptable too, as in $\mathcal{F}_1$. All in all, $\mathcal{F}_3$ has two $<$-preferred extensions $\{R_1, R_2, R_5\}$ and $\{R_1, R_3, R_5, R_8\}$, the latter with conclusions:

- $Cn(\{R_1, R_3, R_5, R_8\}) = \{presentBreast\,Tumour,\ highFatigue,\ elevatedBlood\,Pressure,$

    $R_1, R_3, R_5, R_8,\ Chemotherapy,\ Lymph\,node\,Excision,\ Low\,Int.\,Exercise,$

    $\mathsf{not}\ High\,Int.\,Aerobic\,Exercise,$

    $decreaseBreast\,Tumour,\ increaseFatigue,\ increaseFitness,\ increasePain,$

    $increaseLymphoedema\,Risk,\ decreaseFatigue,\ decreaseFitness,\ decreasePain,$

    $\mathsf{not}\ increaseBlood\,Pressure,$

    $residualBreast\,Tumour,\ normalFatigue,\ \mathsf{not}\ highBlood\,Pressure,$

    $\overline{R_2}, \overline{R_4}, \overline{\mathsf{inapp}(R_1)}, \overline{\mathsf{inapp}(R_2)}, \overline{\mathsf{inapp}(R_3)}, \overline{\mathsf{inapp}(R_5)}, \overline{\mathsf{inapp}(R_8)},$

    $\mathsf{inapp}(R_4), \mathsf{inapp}(R_6), \mathsf{inapp}(R_7), \mathsf{inapp}(R_9), \overline{\mathsf{needs\_repair}(R_1)}\}.$

The induced goal extensions are as follows.

- $\mathcal{G}_{\{R_1,R_2,R_5\}} = \{decreaseBreast\,Tumour, decreaseFatigue, decreaseFitness, decreasePain\};$
- $\mathcal{G}_{\{R_1,R_3,R_5,R_8\}} = \{decreaseBreast\,Tumour, decreaseFatigue, decreaseFitness, decreasePain,$
    $\mathsf{not}\ increaseBlood\,Pressure\}.$

Note that, in accordance with Corollary 4.2, the goal extensions are induced by interaction-free sets of recommendations.

Now, using priorities over goals we find that $\mathsf{not}\ increaseBlood\,Pressure \in \mathcal{G}_{\{R_1,R_3,R_5,R_8\}} \setminus \mathcal{G}_{\{R_1,R_2,R_5\}}$ is trivially such that $g \preccurlyeq_4 \mathsf{not}\ increaseBlood\,Pressure$ for any goal $g \in \mathcal{G}_{\{R_1,R_2,R_5\}} \setminus \mathcal{G}_{\{R_1,R_3,R_5,R_8\}} = \emptyset$. That is, $\mathcal{G}_{\{R_1,R_2,R_5\}} \trianglelefteq_{\mathbb{G}} \mathcal{G}_{\{R_1,R_3,R_5,R_8\}}$. Since $\mathcal{G}_{\{R_1,R_2,R_5\}} \setminus \mathcal{G}_{\{R_1,R_3,R_5,R_8\}} = \emptyset$, it follows that $\mathcal{G}_{\{R_1,R_2,R_5\}} \triangleleft_{\mathbb{G}} \mathcal{G}_{\{R_1,R_3,R_5,R_8\}}$. Therefore, the suggestion by ABA$^+$G for Patient 3 would be to undergo chemotherapy, lymph node excision and take upon low intensity exercise, as well as explicitly to not engage in high intensity aerobic exercise in order not to worsen blood pressure problems.

5.4.4. Patient 4

$\mathcal{F}_4$ is just as $\mathcal{F}_3$, but due to the patient's preference of low intensity exercise over standard exercise, expressed as $R_2 <_4 R_3$, we find $\{R_3\} \rightsquigarrow_< \{R_2\}$ and $\{R_2\} \not\rightsquigarrow_< \{R_3\}$. This renders R_2 not acceptable, so that

$\mathcal{F}_4$ has a unique $<$-preferred extension $\{R_1, R_3, R_5, R_8\}$, with conclusions as in Section 5.4.3. There is thus a unique goal extension $\mathcal{G}_{\{R_1, R_3, R_5, R_8\}}$, induced by this $<$-preferred extension, which is (trivially) the top goal extension and coincides with the one in Section 5.4.3. Observe as well that $\mathcal{G}_{\{R_1, R_3, R_5, R_8\}}$ is induced by the interaction-resolving set of the most preferred applicable recommendations, in accordance with Corollary 4.4.

These illustrations show how ABA$^+$G enables reasoning with interacting clinical guideline recommendation in a patient-centric way, whereby the patient's state and preferences, as well as any goals, set and prioritised by the clinician, are taken into account to provide suggestions as to which recommendations to follow among the many applicable ones. All the $<$-preferred extensions of each ABA$^+$G patient framework $\mathcal{F}_i$, for $i \in \{1, 2, 3, 4\}$, are interaction-free as sets of recommendations that respect the patient's preferences (cf. Theorem 4.1), and the induced (top) goal extensions allow to achieve the desired goals, in accordance with the Ariadne principles (cf. Section 2).

6. Discussion

We here discuss some of our design choices concerning ABA$^+$G in this paper. This discussion pertains to some limitations as well as the potential of the integration of TMR, patient context and ABA$^+$G. A discussion regarding the more general properties of the ABA$^+$G formalism as an extension of ABA$^+$ (e.g. in terms of goal orderings and computational complexity) is omitted because it is beyond the scope of this application-targeted paper.

6.1. Choices in contrast to [30]

We here discuss two TMR-to-ABA$^+$G mapping differences between this paper and [30].

On Definition 3.2. Afforded by TMR, one can in principle identify an interaction's *modal strength* μ, which reflects the conclusiveness of the interaction, and make use of it when reasoning with interacting recommendations. For instance, following the exposition in [108], one could assume, as in [30], that the interaction's modal strength can take two values $\square$ and $\lozenge$, where $\square$ means 'the interaction will certainly occur if the related recommendations are prescribed' and $\lozenge$ means 'the interaction is uncertain to happen' [108]. In principle then, interactions could also be rendered defeasible in ABA$^+$G so as to allow one to argue about the interactions themselves. However, an interaction's modal strength depends on a number of parameters, including the knowledge about the hierarchy of actions (see footnote 10), which is not available in the TMR implementation. We are therefore leaving the defeasibility of interactions for future work.

On Section 4.3.1. In contrast to [30], in Section 4.3.1 we omit an additional condition to rule **12.**, namely *normalGastrointestinal Bleeding*, pertaining to the negative contribution of R_2. That is, instead of rule $\overline{R_1} \leftarrow R_2, normalGastrointestinal\ Bleeding$ (as in [30]) we have $\overline{R_1} \leftarrow R_2$. This change is inconsequential because of the assumptions for the presumed non-applicability of recommendations and rules **8.** and **10.**: a) if *normalGastrointestinal Bleeding* can be argued for, then R_2 cannot be argued to be inapplicable ($\overline{\mathsf{inapp}(R_2)} \leftarrow normalGastrointestinal\ Bleeding$), and so is applicable and can be used (via ($\overline{R_1} \leftarrow R_2$)) to argue against R_1, just like it could be having the rule $\overline{R_1} \leftarrow R_2, normalGastrointestinal\ Bleeding$; b) otherwise, if *normalGastrointestinal Bleeding* cannot be argued for, then R_2 is by default deemed inapplicable ($\overline{R_2} \leftarrow \mathsf{inapp}(R_2)$) and will not be of use in arguing against R_1, just like it would not be having the rule $\overline{R_1} \leftarrow R_2, normalGastrointestinal\ Bleeding$.

6.2. Nuances of ABA$^+$G frameworks

6.2.1. Default applicability of recommendations

We have used predicate inapp() to reason in ABA$^+$G patient frameworks about the applicability of recommendations. Specifically, we effectively declare a recommendation to be applicable, and hence potentially acceptable, when the given patient presents with a property that is affected by the recommendation. In practice, however, as envisaged within the decision support system (DSS) in development under the ROAD2H project, filtering of inapplicable recommendations will be executed outside of the argumentation layer. This means that recommendations (and interactions thereof) provided as TMRweb output for ABA$^+$G to deal with in the context of a patient will all be known to be applicable in advance, because matching the patient's EHR with the TMR ontology will be done by a separate interface. Hence, we could in principle forego constructing assumptions of the form inapp(R) and any rules involving them or their contraries, thus simplifying the mapping from $\mathbb{R}$, $\mathbb{I}$ and $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$ to ABA$^+$G. We have nevertheless shown how ABA$^+$G can fully account for the potential non-applicability of recommendations making it independent of such an interface within the DSS.

6.2.2. Repairable interactions

On Theorem 4.1 and Corollary 4.2. Regarding the way ABA$^+$G resolves interactions among recommendations by yielding $<$-preferred extensions that are interaction-free as sets of recommendations, we note that, in principle, Theorem 4.1 and Corollary 4.2 *cannot* be strengthened to yield interaction-resolving sets of recommendations, as witnessed in the following fictitious example.

Example 6.1. Consider recommendations $\mathbb{R} = \{R_1, R_2\}$ such that $\mathbb{I} = \{(R_1, R_2, Repair)\}$. (We omit to specify the components of the recommendations for simplicity.) Assume an ABA$^+$G patient framework $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$ in which R_1 is applicable, but R_2 is not. So $\mathcal{R}$ contains rules $\overline{R_1} \leftarrow$ inapp(R_1), $\overline{R_2} \leftarrow$ inapp(R_2), $\overline{\text{inapp}(R_1)} \leftarrow normalCondition$ and $normalCondition \leftarrow \top$ for some patient's state $normalCondition \in \mathcal{S}$, as well as $R_2 \leftarrow R_1$, needs_repair(R_1) and $\overline{\text{needs\_repair}(R_1)} \leftarrow R_2$.

Consider the set $E_1 = \{R_1\}$. It is closed (because one cannot deduce R_2 without assuming needs_repair(R_1)), $<$-conflict-free (because no recommendation suffices to deduce its own contrary) and $<$-admissible (because R_1 is applicable, so that $\emptyset \vdash \overline{\text{inapp}(R_1)}$ is an argument, which means that E_1 $<$-defends against the only $<$-attack $\{\text{inapp}(R_1)\} \vdash \overline{R_1}$). Note that, on the one hand, $\{R_1, R_2\}$ is not $<$-admissible, because R_2 is inapplicable, so that $\{\text{inapp}(R_2)\} \vdash \overline{R_2}$ cannot be $<$-defended against. On the other hand, $\{R_1, \text{needs\_repair}(R_1)\}$ is not $<$-admissible, because it is not closed, as it deduces R_2. And while $\{R_1, \text{needs\_repair}(R_1), R_2\}$ is closed, it is likewise not $<$-admissible, for the same reason that $\{R_1, R_2\}$ is not. Consequently, E_1 is $\subseteq$-maximally $<$-admissible, i.e. $<$-preferred. However, $E_1 \cap \mathbb{R} = \{R_1\}$ is not interaction-resolving.

Nonetheless, in practice, a situation such as in Example 6.1 could hardly arise. Indeed, in case of a Repairable interaction between two recommendations, they would be either both applicable or both inapplicable in ABA$^+$G, because the identification of a Repairable interaction in TMR pertains to finding a property that is inversely affected by the actions of the recommendations. In this setting, we would find the $<$-preferred extension yielding an interaction-resolving set of recommendations by using the following line of reasoning.

> If $E \cap \mathbb{R}$ is interaction-free but not interaction-resolving, then for all $R_i, R_j \in \mathbb{R}$ such that $(R_i, R_j, Repair) \in \mathbb{I}$ and $R_i \in E \cap \mathbb{R}$ we find $R_j \notin E \cap \mathbb{R}$. Since $R_i \in E$ implies that R_i is applicable, it must be that any such R_j is applicable too. But then $R_j \notin E$ implies that there is

$R_k \in E$ such that $(R_j, R_k, t) \in \mathbb{I}$ with a) either $t \in \{Repet, Alt\}$, or b) $t = Contr$. In case a), R_k should repair R_i too, by the nature of TMR interactions (that are not spelled out in this paper), whence $R_k \notin E$ leads to a contradiction with $R_k \in E$. Similarly, in case b), R_k should contradict R_i too, which leads to a contradiction with $E \cap \mathbb{R}$ being interaction-free. Then needs_repair$(R_i) \in E$. Therefore, by contradiction, $E \cap \mathbb{R}$ is interaction-resolving.

In any event, our focus was not on unpacking the intricate details of TMR in order to delineate the space of its outcomes. We did not want to couple ABA⁺G to TMR too tightly either, so as to be able to accommodate possible changes in TMR interaction detection mechanisms (cf. [109] versus [108]). It would nevertheless be interesting to study in the future which restrictions on TMR outputs would allow to strengthen the results from Section 4.3.3.

On Definition 4.4.   Regarding the modelling of Repairable interactions in ABA⁺G itself, there are other candidate definitions. For instance, one could define $\mathcal{R}_i^2 = \{R_j \leftarrow R_i : (R_i, R_j, Repair) \in \mathbb{I}\}$. But this would lead to problems as follows (example generated from [108, Fig. 5]). Suppose we have three applicable recommendations R_1, R_2 and R_3 such that $(R_1, R_2, Repair)$, $(R_1, R_3, Repair)$, $(R_2, R_3, Alt) \in \mathbb{I}$. Then $R_2 \leftarrow R_1$, $R_3 \leftarrow R_1$, $\overline{R_2} \leftarrow R_3$, $\overline{R_3} \leftarrow R_2 \in \mathcal{R}$. Hence, any set of assumptions containing R_1 would be closed only if it contained both R_2 and R_3. But note that $\{R_1, R_2, R_3\}$ would not be $<$-conflict-free, which means no $<$-preferred extension could contain R_1. That is, R_1 would not be an acceptable recommendation. Instead, with our current Definition 4.1, $\{R_1, R_2\}$ is closed, $<$-conflict-free and $<$-attacks both $\{$needs_repair$(R_1)\}$ and $\{R_3\}$. Similarly for $\{R_1, R_3\}$. In the end, R_1 can be accepted and necessitates accepting either R_2 or R_3, but not both, as desired.

One could also instead define $\mathcal{R}_i^2 = \{\overline{R_i} \leftarrow$ avoid(R_i), $\overline{\text{avoid}(R_i)} \leftarrow R_j : (R_i, R_j, Repair) \in \mathbb{I}\}$ to model that, if R_i is repairable, then it must have some negative contributions (by the nature of TMR), whence it should generally be avoided, unless one also follows R_j that can repair R_i. But that would be awkward conceptually: a) if R_i has only one contribution, which is henceforth negative, then arguably it should not be a recommendation at all, or should be filtered out by e.g. TMR-web, or else we could instead have $\mathcal{R}_{\mathbb{R}} = \{\overline{\text{inapp}(R)} \leftarrow v_I P : (R, A, \delta, \mathcal{C}) \in \mathbb{R}, (P, E, v_I, v_T) \in \mathcal{C}, (P, E, v_I, v_T)$ is overall positive$\}$, so that R_i would be inapplicable; b) if R_i has multiple contributions and R_i is applicable even with the above redefinition of $\mathcal{R}_{\mathbb{R}}$, then it is still arguably not the case that one should 'avoid' R_i. We instead believe and have shown that our current definition works well.

6.2.3. Preferences over recommendations

We lastly discuss briefly our theoretical results concerning preferences over recommendations.

Note that the preference relation over recommendations in Theorem 4.4 and Corollary 4.5 is required to be total. Indeed, with partial preferences the results would not hold in general. This is illustrated in the following example, which is a reformulation of a canonical counter-example [92, Example 4.1] to capturing the so-called *preferred sub-theories* [17] (in the general case of partially ordered defaults without linearisation of preferences) in structured argumentation.

Example 6.2. Consider fictitious recommendations (with some components left unspecified, indicated with '?'):

- $R_a = (R_a, A, \text{should}, (P^A, E^a, ?, ?))$,
- $R_{\neg a} = (R_{\neg a}, A, \text{should not}, (P^A, E^{\neg a}, ?, ?))$,
- $R_b = (R_b, B, \text{should}, (P^B, E^b, ?, ?))$,
- $R_{\neg b} = (R_{\neg b}, B, \text{should not}, (P^B, E^{\neg b}, ?, ?))$.

Thus $\mathbb{R} = \{R_a, R_{\neg a}, R_b, R_{\neg b}\}$ and $\mathbb{I} = \{(R_a, R_{\neg a}, Contr), (R_b, R_{\neg b}, Contr)\}$. Assume that the patient's context $(\mathcal{S}, \mathcal{G}, \leqslant, \preccurlyeq)$ defines state $\mathcal{S} = \{P^A, P^B\}$ as well as partial preferences $R_b < R_a$ and $R_{\neg a} < R_{\neg b}$. Put simply, we have two pairs $(R_a, R_{\neg a})$ and $(R_b, R_{\neg b})$ of mutually contradicting but not preference-related (within the pairs) recommendations that are applicable in the given context (there would be rules such as $\overline{\mathsf{inapp}(R_a)} \leftarrow P^A$ and $P^A \leftarrow \top$ in the resulting ABA$^+$G patient framework $\mathcal{F}_p = (\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, \leqslant, \mathcal{G}, \preccurlyeq)$). The set $\mathbb{R}' = \{R_a, R_{\neg b}\}$ of the most preferred recommendations is interaction-resolving. Then, in $\mathcal{F}_p$, $\{R_{\neg a}\} \rightsquigarrow_< \{R_a\}$ and $\{R_b\} \rightsquigarrow_< \{R_{\neg b}\}$ due to the existence of rules $\overline{R_a} \leftarrow R_{\neg a}$, $\overline{R_{\neg b}} \leftarrow R_b$ and the fact that preferences do not relate the corresponding recommendations. Since all recommendations are applicable, we find that, specifically, $\{R_{\neg a}, R_b\}$ is a $<$-preferred extension. However, $\mathbb{R}' \nsubseteq \{R_{\neg a}, R_b\}$.

It would be interesting to investigate other restrictions on preferences over $\mathbb{R}$ that relax the totality requirement but allow to preserve the same result as in Theorem 4.4. However, note that once partial orders are concerned, one can have two applicable recommendations R_1 and R_2 that are e.g. in Contradiction interaction and are not related by preferences among themselves or with respect to any other recommendations. As such, R_1 and R_2 would be $\leqslant$-maximal applicable, but $\{R_1, R_2\}$ would not be $<$-conflict-free. So the set $\mathbb{R}'$ of the most preferred applicable recommendations would not be interaction-resolving, thus defeating the purpose of analysing whether such recommendations would be contained in every $<$-preferred extension. Nonetheless, we leave it for future work to study variations of Theorem 4.4.

6.3. Differences from other argumentation formalisms

We stated in the Introduction that the main reasons for using ABA$^+$ in this work are its rule-based nature together with preference-enabled, extension-based reasoning mechanisms as well as the availability of implementations of ABA$^+$. The rule-based formalisation allows ABA$^+$G to naturally model recommendations and their components expressed in TMR. The mechanism of dealing with preferences in ABA$^+$ allows ABA$^+$G to express preferences over recommendations and ensure that interactions among them are resolved. The semantics of $<$-preferred extensions allows for credulous choices in ABA$^+$G among interacting recommendations. Very importantly, existing complexity analysis [15,34,60] of (versions of) ABA and ABA$^+$ as well as their implementations [7,56] enable rapid development of ABA$^+$G for deployment in an argumentation-assisted DSS.

Predominantly for these reasons, we chose ABA$^+$ instead of some other prominent argumentation formalisms. For instance, Value-Based Argumentation [8,9,54] could potentially be used to reason about different values brought about by following recommendations. However, since TMR lends itself to be naturally formulated using rules and arguably less so using abstract arguments and values, we chose a structured argumentation formalism appropriate for TMR. Carneades [45,47] could also potentially be used in modelling TMR representations via argumentation schemes, yet this would introduce another layer of formalisation and complexity. It also, along with Defeasible Logic Programming (DeLP) [43], employs a sceptical reasoning mechanism, whereas ABA$^+$G instead follows a credulous semantics to allow for choosing among equally legitimate options in the context of contradiction, alternative and repetition interactions among recommendations.

ASPIC$^+$ [67,68,76] is another natural candidate for reasoning with TMR representations, as posited in [29]. One major difference between the use of ASPIC$^+$ and ABA$^+$ could be their treatment of preferences and, specifically in our setting, asymmetric attacks. Indeed, if more nuanced interactions among recommendations were allowed, as is envisaged in the future, then their knowledge representation would entail that the attacks between, say, two recommendations in contradiction, would not be symmetric as

it is now in ABA$^+$G, but instead depend on the deducibility of other sentences. As a consequence, reversing (as in ABA$^+$) and deleting (as in ASPIC$^+$) attacks would not result into the same conflict-free sets, and consequently the same desirable outcomes.

In depth analysis is beyond the scope of this paper but it would indeed be interesting to see in the future whether the reasoning as in ABA$^+$G could be enabled by other formalisms, including those discussed above and e.g. Deductive Argumentation [12,13,48] and DefLog [97,98].

We have argued for certain desirable characteristics of our mapping from TMR and patient context into ABA$^+$G, and illustrated that it is applicable in practice. It would nonetheless be interesting to study in the future its possible improvements conceptually and reasoning-wise, especially the currently unaddressed aspects of TMR.

7. Related work

Argumentation (with or without preferences) has been successfully applied in health care, see e.g. [6,63] for overviews. We discuss several strands of research in this setting with notable examples.

7.1. Argumentation for medical reasoning in general

Several works use argumentation semantics for reasoning with medical knowledge and employing preferences. For instance, in [52], manually extracted evidence from randomised clinical trials and systematic reviews is 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 [35] are used to discard certain attacks, whence semantics of grounded and preferred extensions are used to identify the acceptable arguments. The focus is determining superiority among treatments, not concerning guideline recommendations or conflict resolution among those.

Other works, e.g. [40,78,94], integrate argumentation with preferences to help clinicians to construct, exchange and evaluate arguments for and against decisions. For instance, in [94] argumentation with its semantics as well as preferences are used in multi-agent setting deliberation about organ transplantation. Experts use argumentation schemes [46,102] to construct arguments and attacks concerning viability of transplantation. A mediator agent evaluates the arguments by determining their strength using guideline knowledge as preferences, knowledge about past transplantation as well as about the interacting agents. Similar in spirit is the system ArgMed [78] that allows to document and turn clinicians' discussions into argumentation frameworks using argumentation schemes, whence preferred semantics is used to find the best claims. On the other hand, to automate medical reasoning, in [40], agents are proposed to exchange arguments structured with claims and backings, and thus arguable against via the latter, but instead of (classical) argumentation semantics various argument weighing and aggregation mechanisms are intended to be used to support decision making. In any event, these works do not concern reasoning with guidelines.

7.2. Argumentation for reasoning with clinical guidelines

Argument aggregation for reasoning with guidelines is used in e.g. [50,106]. Specifically, in [50] argumentation schemes are used as templates for generating arguments that correspond to statements in guidelines. In particular, an argument consists of assumptions, claim, polarity (for or against the

claim), confidence (e.g. quality of evidence, likelihood of an outcome) and precondition (i.e. whether the argument is applicable). To perform reasoning, a single goal must be specified, whence confidence of arguments is aggregated to identify the acceptable arguments so as to achieve that goal. Similarly, [106] employs a form of argumentation to weigh and aggregate arguments for and against candidate decisions constructed from guidelines towards achieving specified goals. The focus of these works is enacting recommendations from a single guideline, rather than reasoning with interacting recommendations from multiple guidelines.

In terms of reasoning with multiple guidelines in the setting of multimorbidities, the recent CONSULT project [19,20,59,107] applies argumentation to reason with guidelines and patient preferences for managing patients with comorbidities. Specifically, they use structured metalevel argumentation frameworks (MAFs) based on either, essentially, second-order logic as in [107], or first-order logic as in [59], to construct arguments using argumentation schemes and particularly the *critical questions* pertaining to the latter. Their newly introduced argumentation schemes with their critical questions serve as templates for structuring arguments about statements manually extracted from clinical guidelines. They further integrate preferences modelled as attacks on attacks, following [65], into MAFs to resolve conflicts among arguments. What is more, they also consider conflicts among guideline recommendations by means of TMR as an external service to the argumentation engine [19]. Importantly, the CONSULT project has developed and uses TMRweb. We likewise rely on TMRweb in our implementation efforts, but in our theoretical foundations instead use the TMR model to represent guideline recommendations and identify their interactions, which we then together with patient information map directly into ABA$^+$G frameworks (using effectively a fragment of Horn logic contrasting greatly complexity-wise with first- and second-order logics). We also incorporate preference information directly in the construction of attacks when resolving interactions. In addition, we allow for reasoning with prioritised goals in ABA$^+$G. Importantly, our approach enables us to meet the Ariadne principles of patient management.

7.3. Non-argumentative approaches to medical decision making

Non-argumentative approaches to reasoning with clinical guidelines exist too, see [75,82] for overviews. A recent work concerning reasoning with interacting guidelines, patient conditions and preferences represents guideline recommendations as actionable graphs [105], mapping them into first-order logic (FOL) rules, while representing patient conditions and preferences as FOL revision operators. Then, 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 – the 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 focus on execution of single guidelines, e.g. [61,89], or identification of incompatibilities among guidelines. As to the latter, answer set programming is for instance used in [91] to check temporal conformance through *a posteriori* verification of a single guideline with the recommendations actually followed, motivated by the patient state. On the other hand, statistical preference learning is used in [96] to identify inconsistencies in antibiotic therapy guidelines. The objectives of these works are thus different from the objective of the work herein.

Yet other works concern preference elicitation to facilitate clinical decision making. In particular, in [83] the authors incorporate patients' preferences in terms of QALY (quality-adjusted life-year), utilities

and costs into the shared decision making model. 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 nonetheless to see how this line of work could inform knowledge representation in our approach.

7.4. Goal-driven argumentative decision making

Goal-driven argumentative decision making (possibly with preferences) has been explored, see e.g. [2,36,71,110]. For instance, the approach of [2] concerns general multiple criteria decision making in argumentation with preferences via reasoning backwards from goals to arguments. A follow-up application-specific approach (see [71]) affords goal-driven argumentative documentation, analysis and making of decisions. On the one hand, the settings there do not apply to reasoning with guidelines. On the other hand, ABA+G differs from these approaches in at least two other aspects. First, in terms of using preferences (over goals) to select among extensions, as in e.g. [4,100]. Second, in terms of the direction of reasoning – from arguments to goals, which is more similar to assumption-based reasoning with goals and preferences as in [36], which we discuss in more detail below.

In [36] ABA frameworks without preferences are used for contract negotiation. The frameworks are however equipped with goals and preferences among them. Therein, a goal corresponds to the conclusion of a rule and preferences over goals follow a total preorder. These ideas are also adopted in our work, along with the reasoning principle of pursuing higher-ranked goals at the expense of lower-ranked goals. Thus, comparing solutions in the form of $<$-preferred extensions amounts to comparing goal states in [36], which correspond to goal extensions in our work. The approaches are however slightly different in details. For comparison, we reproduce here their definition of an ordering of goal states:

For G and G' goal states, G is preferred to G', denoted by $G \sqsupseteq G'$, iff

(1) there exists a goal g that is satisfied in G but not in G', and
(2) for each goal g', if $P(g') \geqslant P(g)$ and g' is satisfied in s', then g' is also satisfied in s,

where P is a ranking function mapping goals to natural numbers.

Now, for contrast, let G and G' be two goal states such that $G \setminus \{g\} = G' \setminus \{g'\}$ and $g \neq g'$ are equally preferred, i.e. $P(g) = P(g')$. By the above definition it is not the case that $G \sqsupseteq G'$ nor it is the case that $G' \sqsupseteq G$. Thus, G and G' are incomparable with respect to $\sqsupseteq$. More generally, goal states differing only in goals that are equally preferred are incomparable. Yet, the authors in [36] then declare such goal states to be 'equally preferred', albeit without defining this notion with respect to $\sqsupseteq$. Instead, according to our Definition 4.3, we find $G \trianglelefteq_{\mathbb{G}} G'$ and $G' \trianglelefteq_{\mathbb{G}} G$, so that G and G' are comparable and equally preferred with respect to $\trianglelefteq_{\mathbb{G}}$. (Note that, in general, $G \sqsupseteq G'$ implies $G' \trianglelefteq_{\mathbb{G}} G$ by taking a $\preccurlyeq$-maximal g (i.e. with $\leqslant$-maximal $P(g)$); the converse does not hold, however, because Definition 4.3 does not concern the common goals $G \cap G'$.)

In [36], the authors further elaborate on goal states by proposing minimal goal states from a set of states promoted by a decision. Therein, a decision is an accepted assumption and all assumptions representing decisions are mutually exclusive. The possible minimal states stemming from decisions are used to characterise single assumptions and establish preferences among them. In ABA+G instead, assumptions representing recommendations are not generally mutually exclusive, with preferences among them

used for establishing acceptable recommendations and goals thereof. Overall thus, the objectives of our work and that of [36] are rather different. It would nonetheless be interesting to investigate the formal relationships with this work in the future.

In relation to argumentative decision making, we also mention the general approach of [2], where several principles for the comparison of decisions are established. The aims of that work and ours are significantly different, however. The authors of [2] analyse abstract argumentation for the purpose of general decision-making and use pre-established candidate decisions within the argumentative reasoning. Differently, here we deploy structured argumentation in the form of ABA$^+$ and augment it with goals to accommodate the TMR model and meet the Ariadne principles for a domain-specific application, whereby decisions are formed after the argumentative reasoning with extension-based semantics.

We lastly note that an argumentative approach to explainable decision making with contextual goals was recently proposed and illustrated with a medical decision making example in [110]. There, context rules and primitives involving patient state properties are used to assert defeasibility of logical implications between decisions, attributes, and goals. While the approach focuses on *explainability* issues in decision making and is thus not directly related to this paper, it shows that context-sensitivity is an important and desirable property in both medical and argumentative settings, that we specifically addressed in this work.

8. Conclusions and future work

We have shown how ABA$^+$G, a structured-argumentation formalism proposed in [30] and extending the ABA$^+$ [15,27,33] formalism with prioritised goals, can be used to automate patient-centric reasoning with interacting clinical guideline recommendations. Specifically, we mapped Transition-based Medical Recommendation (TMR) [108,109] representations of guideline recommendations to ABA$^+$, incorporated in ABA$^+$ patient-specific conditions and their preferences, and augmented ABA$^+$ to ABA$^+$G so as to deal with patient-centric goals and priorities among them. We showed, among other properties, that ABA$^+$G yields interaction-free sets of recommendations taking into account the context of the patient in terms of their state, preferences (over recommendation actions), and prioritised achievable goals. We illustrated our approach to patient-centric reasoning with interacting guideline recommendations by using a TMR-based use case, complementing it with various patients and their contexts. We posited that our approach meets the set-out Ariadne principles [72] of patient management, thus establishing a unique relationship between features of argumentative reasoning and personalised care in multimorbidity settings.

The most important milestone in the future is carrying out an evaluation of ABA$^+$G as well as integrating it within an overall decision support system (DSS) assisting with decision making in multimorbidity settings, for instance as envisaged within the ROAD2H project[21] in real clinical settings. A crucial aspect of such an evaluation will be the explainability of the overall system. Argumentation is indeed well-suited for explainable reasoning [6,40,70] with argumentative explanations proposed in various settings, see e.g. [1–3,5,14,16,18,21–24,26,28,31,32,37–39,42,43,51,57,58,62,64,66,69,70,77,79,81,84–88,90,93,101,103,104,110–112]. We hope to exploit the well-established as well as novel ABA$^+$ mechanisms to our advantage of providing various explanations to accompany the decisions supported by ABA$^+$G. In addition to several other future work directions mentioned in Sections 6 and 7, we will aim

---

[21]www.road2h.org

to extend ABA$^+$G to take into account various TMR artefacts not yet present in ABA$^+$G. This may yield additional preferences and result in a probabilistic extension of ABA$^+$G, requiring further study.

Acknowledgements

Data access statement

No new data was collected in the course of this research.

References

[1] E. Albini, A. Rago, P. Baroni and F. Toni, Relation-based counterfactual explanations for Bayesian network classifiers, in: *29th International Joint Conference on Artificial Intelligence*, C. Bessiere, ed., IJCAI, Yokohama, 2020, pp. 451–457. doi:10.24963/ijcai.2020/63.

[2] L. Amgoud and H. Prade, Using arguments for making and explaining decisions, *Artificial Intelligence* **173**(3–4) (2009), 413–436. ISSN 0004-3702. doi:10.1016/j.artint.2008.11.006.

[3] L. Amgoud and M. Serrurier, Agents that argue and explain classifications, *Autonomous Agents and Multi-Agent Systems* **16**(2) (2008), 187–209. doi:10.1007/s10458-007-9025-6.

[4] L. Amgoud and S. Vesic, Rich preference-based argumentation frameworks, *International Journal of Approximate Reasoning* **55**(2) (2014), 585–606. doi:10.1016/j.ijar.2013.10.010.

[5] A. Arioua, N. Tamani and M. Croitoru, Query answering explanation in inconsistent datalog +/− knowledge bases, in: *Database and Expert Systems Applications – 26th International Conference*, Q. Chen, A. Hameurlain, F. Toumani, R. Wagner and H. Decker, eds, Lecture Notes in Computer Science, Vol. 9261, Springer, Valencia, 2015, pp. 203–219, ISSN 16113349. doi:10.1007/978-3-319-22849-5_15.

[6] K. Atkinson, P. Baroni, M. Giacomin, A. Hunter, H. Prakken, C. Reed, G.R. Simari, M. Thimm and S. Villata, Towards artificial argumentation, *AI Magazine* **38**(3) (2017), 25–36. doi:10.1609/aimag.v38i3.2704.

[7] Z. Bao, K. Čyras and F. Toni, ABAplus: Attack reversal in abstract and structured argumentation with preferences, in: *Principles and Practice of Multi-Agent Systems – 20th International Conference*, B. An, A.L.C. Bazzan, J. Leite, S. Villata and L. van der Torre, eds, Lecture Notes in Computer Science, Springer, Nice, 2017, pp. 420–437. doi:10.1007/978-3-319-69131-2_25.

[8] T.J.M. Bench-Capon, Persuasion in practical argument using value-based argumentation frameworks, *Journal of Logic and Computation* **13**(3) (2003), 429–448. doi:10.1093/logcom/13.3.429.

[9] T.J.M. Bench-Capon, K. Atkinson and A. Chorley, Persuasion and value in legal argument, *Journal of Logic and Computation* **15**(6) (2005), 1075–1097. ISSN 0955-792X. doi:10.1093/logcom/exi058.

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

[11] P. Besnard, A.J. García, A. Hunter, S. Modgil, H. Prakken, G.R. Simari and F. Toni, Introduction to structured argumentation, *Argument & Computation* **5**(1) (2014), 1–4. doi:10.1080/19462166.2013.869764.

[12] P. Besnard and A. Hunter, A logic-based theory of deductive arguments, *Artificial Intelligence* **128**(1–2) (2001), 203–235. doi:doi:10.1016/S0004-3702(01)00071-6.

[13] P. Besnard and A. Hunter, Constructing argument graphs with deductive arguments: A tutorial, *Argument & Computation* **5**(1) (2014), 5–30. doi:10.1080/19462166.2013.869765.

[14] F. Bex and D. Walton, Combining explanation and argumentation in dialogue, *Argument & Computation* **7**(1) (2016), 55–68. doi:10.3233/AAC-160001.

[15] A. Bondarenko, P.M. Dung, R. Kowalski and F. Toni, An abstract, argumentation-theoretic approach to default reasoning, *Artificial Intelligence* **93**(97) (1997), 63–101. doi:10.1016/S0004-3702(97)00015-5.

[16] R. Booth, D.M. Gabbay, S. Kaci, T. Rienstra and L. van der Torre, Abduction and dialogical proof in argumentation and logic programming, in: *21st European Conference on Artificial Intelligence*, T. Schaub, G. Friedrich and B. O'Sullivan, eds, Frontiers in Artificial Intelligence and Applications, Vol. 263, IOS Press, Prague, 2014, pp. 117–122. ISBN 9781614994190. doi:10.3233/978-1-61499-419-0-117.

[17] G. Brewka, Preferred subtheories: An extended logical framework for default reasoning, in: *11th International Joint Conference on Artificial Intelligence*, N.S. Sridharan, ed., Morgan Kaufmann, Detroit, 1989, pp. 1043–1048.

[18] C.E. Briguez, M.C.D. Budán, C.A.D. Deagustini, A.G. Maguitman, M. Capobianco and G.R. Simari, Argument-based mixed recommenders and their application to movie suggestion, *Expert Systems with Applications* **41**(14) (2014), 6467–6482. doi:10.1016/j.eswa.2014.03.046.

[19] M. Chapman, P. Balatsoukas, M. Ashworth, V. Curcin, N. Kökciyan, N. Essers, I. Sassoon, S. Modgil, S. Parsons and E.I. Sklar, Computational argumentation-based clinical decision support demonstration, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, N. Agmon, E. Elkind, M.E. Taylor and M. Veloso, eds, IFAAMAS, Montreal, 2019, pp. 2345–2347.

[20] M.D. Chapman and V. Curcin, A microservice architecture for the design of computer-interpretable guideline processing tools, in: *18th IEEE International Conference on Smart Technologies*, IEEE Computer Society Press, Novi Sad, 2019.

[21] O. Cocarascu, A. Rago and F. Toni, Extracting dialogical explanations for review aggregations with argumentative dialogical agents, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, E. Elkind, M. Veloso, N. Agmon and M.E. Taylor, eds, IFAAMAS, Montreal, 2019, pp. 1261–1269, ISSN 15582914. ISBN 9781510892002.

[22] O. Cocarascu, A. Stylianou, K. Cyras and F. Toni, Data-empowered argumentation for dialectically explainable predictions, in: *24th European Conference on Artificial Intelligence*, G.D. Giacomo, A. Catalá, B. Dilkina, M. Milano, S. Barro, A. Bugarín and J. Lang, eds, IOS Press, Santiago de Compostela, 2020, pp. 2449–2456. doi:10.3233/FAIA200377.

[23] J. Collenette, K. Atkinson and T.J.M. Bench-Capon, An explainable approach to deducing outcomes in European court of human rights cases using ADFs, in: *Computational Models of Argument*, H. Prakken, ed., IOS Press, 2020, pp. 21–32. doi:10.3233/FAIA200488.

[24] A. Collins, D. Magazzeni and S. Parsons, Towards an argumentation-based approach to explainable planning, in: *2nd International Workshop on Explainable AI Planning*, T. Chakraborti, D. Dannenhauer, J. Hoffmann and D. Magazzeni, eds, Berkeley, CA, 2019.

[25] K. Čyras, ABA+: Assumption-based argumentation with preferences, PhD thesis, Imperial College London, 2017, https://spiral.imperial.ac.uk/handle/10044/1/58340.

[26] K. Čyras, D. Birch, Y. Guo, F. Toni, R. Dulay, S. Turvey, D. Greenberg and T. Hapuarachchi, Explanations by arbitrated argumentative dispute, *Expert Systems with Applications* **127** (2019), 141–156. doi:10.1016/j.eswa.2019.03.012.

[27] K. Čyras, X. Fan, C. Schulz and F. Toni, Assumption-based argumentation: Disputes, explanations, preferences, in: *Handbook of Formal Argumentation*, Vol. 1, P. Baroni, D.M. Gabbay, M. Giacomin and L. van der Torre, eds, College Publications, 2018. ISBN 9781848902756.

[28] K. Čyras, D. Letsios, R. Misener and F. Toni, Argumentation for explainable scheduling, in: *33rd AAAI Conference on Artificial Intelligence*, AAAI Press, Honolulu, Hawaii, 2019, pp. 2752–2759.

[29] K. Čyras and T. Oliveira, Argumentation for reasoning with conflicting clinical guidelines and preferences, in: *Principles of Knowledge Representation and Reasoning, 16th International Conference*, M. Thielscher, F. Toni and F. Wolter, eds, AAAI Press, Tempe, AZ, 2018, pp. 631–632.

[30] K. Čyras and T. Oliveira, Resolving conflicts in clinical guidelines using argumentation, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, N. Agmon, E. Elkind, M.E. Taylor and M. Veloso, eds, IFAAMAS, Montreal, 2019, pp. 1731–1739.

[31] K. Čyras, K. Satoh and F. Toni, Abstract argumentation for case-based reasoning, in: *15th International Conference on Principles of Knowledge Representation and Reasoning*, C. Baral, J.P. Delgrande and F. Wolter, eds, AAAI Press, Cape Town, 2016, pp. 549–552.

[32] K. Čyras, K. Satoh and F. Toni, Explanation for case-based reasoning via abstract argumentation, in: *6th International Conference on Computational Models of Argument*, IOS Press, Potsdam, 2016, pp. 243–254.

[33] K. Čyras and F. Toni, ABA+: Assumption-based argumentation with preferences, in: *Principles of Knowledge Representation and Reasoning, 15th International Conference*, C. Baral, J.P. Delgrande and F. Wolter, eds, AAAI Press, Cape Town, 2016, pp. 553–556.

[34] Y. Dimopoulos, B. Nebel and F. Toni, On the computational complexity of assumption-based argumentation for default reasoning, *Artificial Intelligence* **141**(1–2) (2002), 57–78. doi:10.1016/S0004-3702(02)00245-X.

[35] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artificial Intelligence* **77**(2) (1995), 321–357. doi:10.1016/0004-3702(94)00041-X.

[36] P.M. Dung, P.M. Thang and F. Toni, Towards argumentation-based contract negotiation, in: *Computational Models of Argument*, P. Besnard, S. Doutre and A. Hunter, eds, Frontiers in Artificial Intelligence and Applications, Vol. 172, IOS Press, Toulouse, 2008, pp. 134–146. ISBN 978-1-58603-859-5.

[37] X. Fan, On generating explainable plans with assumption-based argumentation, in: *21th International Conference on Principles and Practice of Multi-Agent Systems*, T. Miller, N. Oren, Y. Sakurai, I. Noda, B.T.R. Savarimuthu and T. Cao Son, eds, Springer, Cham, 2018, pp. 344–361. ISBN 978-3-030-03098-8. doi:10.1007/978-3-030-03098-8_21.

[38] X. Fan and F. Toni, On computing explanations in argumentation, in: *29th AAAI Conference on Artificial Intelligence*, B. Bonet and S. Koenig, eds, AAAI Press, Austin, Texas, 2015, pp. 1496–1502. ISBN 978-1-57735-698-1.

[39] X. Fan and F. Toni, On computing explanations for non-acceptable arguments, in: *Theory and Applications of Formal Argumentation – 3rd International Workshop*, E. Black, S. Modgil and N. Oren, eds, Lecture Notes in Computer Science, Vol. 9524, Springer, Buenos Aires, 2015, pp. 112–127. doi:10.1007/978-3-319-28460-6_7.

[40] J. Fox, L. Black, D. Glasspool, S. Modgil, A. Oettinger, V. Patkar and M. Williams, Towards a general model for argumentation services, in: *Argumentation for Consumers of Healthcare, Papers from the 2006 AAAI, Spring Symposium*, AAAI, Stanford, CA, 2006, pp. 52–57.

[41] P. Fraccaro, M. Arguello Castelerio, J. Ainsworth and I. Buchan, Adoption of clinical decision support in multimorbidity: A systematic review, *JMIR Medical Informatics* **3**(1) (2015). ISSN 2291-9694. doi:10.2196/medinform.3503.

[42] A.J. García, C. Chesñevar, N. Rotstein and G.R. Simari, Formalizing dialectical explanation support for argument-based reasoning in knowledge-based systems, *Expert Systems with Applications* **40** (2013), 3233–3247. doi:10.1016/j.eswa.2012.12.036.

[43] A.J. García and G.R. Simari, Defeasible logic programming: DeLP-servers, contextual queries, and explanations for answers, *Argument & Computation* **5**(1) (2014), 63–88. doi:10.1080/19462166.2013.869767.

[44] GOLD, Global strategy for the diagnosis, management and prevention of COPD, global initiative for chronic obstructive lung disease, https://goldcopd.org/gold-2017-global-strategy-diagnosis-management-prevention-copd/, Accessed: 2019-08-15.

[45] T.F. Gordon, H. Prakken and D. Walton, The Carneades model of argument and burden of proof, *Artificial Intelligence* **171**(10–15) (2007), 875–896. ISSN 0004-3702. doi:10.1016/j.artint.2007.04.010.

[46] T.F. Gordon and D. Walton, Legal reasoning with argumentation schemes, in: *12h International Conference on Artificial Intelligence and Law*, ACM, Barcelona, 2009, pp. 137–146. ISBN 9781605585970. doi:10.1145/1568234.1568250.

[47] T.F. Gordon and D. Walton, Formalizing balancing arguments, in: *Computational Models of Argument*, P. Baroni, T.F. Gordon, T. Scheffler and M. Stede, eds, Frontiers in Artificial Intelligence and Applications, Vol. 287, IOS Press, Potsdam, 2016, pp. 327–338. ISBN 9781614996866. doi:10.3233/978-1-61499-686-6-327.

[48] N. Gorogiannis and A. Hunter, Instantiating abstract argumentation with classical logic arguments: Postulates and properties, *Artificial Intelligence* **175**(9–10) (2011), 1479–1497. doi:10.1016/j.artint.2010.12.003.

[49] A. Grace, C. Mahony, J. O'Donoghue, T. Heffernan, D. Molony and T. Carroll, Evaluating the effectiveness of clinical decision support systems: The case of multimorbidity care, *Journal of Decision Systems* **22**(2) (2013), 97–108. doi:10.1080/12460125.2013.780320.

[50] M.A. Grando, D. Glasspool and A.A. Boxwala, Argumentation logic for the flexible enactment of goal-based medical guidelines, *Journal of Biomedical Informatics* **45**(5) (2012), 938–949. ISSN 1532-0480 (Electronic) 1532-0464 (Linking). doi:10.1016/j.jbi.2012.03.005.

[51] A. Hecham, A. Arioua, G. Stapleton and M. Croitoru, An empirical evaluation of argumentation in explaining inconsistency-tolerant query answering, in: *30th International Workshop on Description Logics*, A. Artale, B. Glimm and R. Kontchakov, eds, CEUR-WS.org, Montpellier, 2017, ISSN 16130073.

[52] A. Hunter and M. Williams, Aggregating evidence about the positive and negative effects of treatments, *Artificial Intelligence in Medicine* **56**(3) (2012), 173–190. ISSN 0933-3657. doi:10.1016/j.artmed.2012.09.004.

[53] S. Kaci and N. Patel, A postulate-based analysis of comparative preference statements, *Journal of Applied Logic* **12**(4) (2014), 501–521. ISSN 1570-8683. doi:10.1016/j.jal.2014.07.004.

[54] S. Kaci and L. van der Torre, Preference-based argumentation: Arguments supporting multiple values, *International Journal of Approximate Reasoning* **48**(3) (2008), 730–751. doi:10.1016/j.ijar.2007.07.005.

[55] A.C. Kakas and P. Moraitis, Argumentation based decision making for autonomous agents, in: *2nd International Joint Conference on Autonomous Agents & Multiagent Systems*, ACM Press, Melbourne, 2003, pp. 883–890. doi:10.1145/860575.860717.

[56] A. Karamlou, K. Čyras and F. Toni, Complexity results and algorithms for bipolar argumentation, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, N. Agmon, E. Elkind, M.E. Taylor and M. Veloso, eds, IFAAMAS, Montreal, 2019, pp. 1713–1721, http://www.ifaamas.org/Proceedings/aamas2019/pdfs/p1713.pdf, http://dl.acm.org/citation.cfm?id=3331902.

[57] A. Karamlou, K. Čyras and F. Toni, Deciding the winner of a debate using bipolar argumentation, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, N. Agmon, E. Elkind, M.E. Taylor and M. Veloso, eds, IFAAMAS, Montreal, 2019, pp. 2366–2368.

[58] N. Kökciyan, S. Parsons, I. Sassoon, E. Sklar and S. Modgil, An argumentation-based approach to generate domain-specific explanations, in: *European Conference on Multiagent Systems*, 2020.

[59] N. Kökciyan, I. Sassoon, A.P. Young, M. Chapman, T. Porat, M. Ashworth, V. Curcin, S. Modgil, S. Parsons and E. Sklar, Towards an argumentation system for supporting patients in self-managing their chronic conditions, in: *Joint Workshop on Health Intelligence (W3PHIAI)*, New Orleans, Louisiana, 2018.

[60] T. Lehtonen, J.P. Wallner and M. Järvisalo, Reasoning over assumption-based argumentation frameworks via direct answer set programming encodings, in: *33rd AAAI Conference on Artificial Intelligence*, AAAI Press, Honolulu, Hawaii, 2019, pp. 2938–2945.

[61] G. Leonardi, A. Bottrighi, G. Galliani, P. Terenziani, A. Messina and F. Della Corte, Exceptions handling within GLARE clinical guideline framework, in: *AMIA Annual Symposium Proceedings*, Vol. 2012, 2012, pp. 512–521, ISSN 1942-597X.

[62] B. Liao and L. van der Torre, Explanation semantics for abstract argumentation, in: *Computational Models of Argument*, Vol. 326, H. Prakken, ed., IOS Press, 2020, pp. 271–282. doi:10.3233/FAIA200511.

[63] L. Longo, 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*, Vol. 9605, A. Holzinger, ed., Springer, 2016, pp. 183–208. doi:10.1007/978-3-319-50478-0_9.

[64] P. Madumal, T. Miller, L. Sonenberg and F. Vetere, A grounded interaction protocol for explainable artificial intelligence, in: *18th International Conference on Autonomous Agents and MultiAgent Systems*, E. Elkind, M. Veloso, N. Agmon and M.E. Taylor, eds, IFAAMAS, Montreal, 2019, pp. 1033–1041.

[65] S. Modgil, Reasoning about preferences in argumentation frameworks, *Artificial Intelligence* **173**(9–10) (2009), 901–934. doi:10.1016/j.artint.2009.02.001.

[66] S. Modgil and M. Caminada, Proof theories and algorithms for abstract argumentation frameworks, in: *Argumentation in Artificial Intelligence*, G.R. Simari and I. Rahwan, eds, Springer, 2009, pp. 105–129, Chapter 6. doi:10.1007/978-0-387-98197-0_6.

[67] S. Modgil and H. Prakken, A general account of argumentation with preferences, *Artificial Intelligence* **195** (2013), 361–397. doi:10.1016/j.artint.2012.10.008.

[68] S. Modgil and H. Prakken, The ASPIC+ framework for structured argumentation: A tutorial, *Argument & Computation* **5**(1) (2014), 31–62. doi:10.1080/19462166.2013.869766.

[69] M. Morveli-Espinoza and C.A. Tacla, Towards an explainable argumentation-based agent, in: *9th European Starting AI Researchers' Symposium*, S. Rudolph and G. Marreiros, eds, CEUR-WS.org, Santiago de Compostela, 2020.

[70] B. Moulin, H. Irandoust, M. Bélanger and G. Desbordes, Explanation and argumentation capabilities: Towards the creation of more persuasive agents, *Artificial Intelligence Review* **17**(3) (2002), 169–222. ISSN 0269-2821. doi:10.1023/A:1015023512975.

[71] J. Muller and A. Hunter, An argumentation-based approach for decision making, in: *24th International Conference on Tools with Artificial Intelligence*, Vol. 1, IEEE, 2012, pp. 564–571, ISSN 1082-3409. ISBN 978-1-4799-0227-9. doi:10.1109/ICTAI.2012.82.

[72] C. Muth, M. van den Akker, J.W. Blom, C.D. Mallen, J. Rochon, F.G. Schellevis, A. Becker, M. Beyer, J. Gensichen, H. Kirchner, R. Perera, A. Prados-Torres, M. Scherer, U. Thiem, H. van den Bussche and P.P. Glasziou, The ariadne principles: How to handle multimorbidity in primary care consultations, *BMC Medicine* **12**(1) (2014), 1–11. ISSN 1741-7015. doi:10.1186/s12916-014-0223-1.

[73] T. Oliveira, J. Dauphin, K. Satoh, S. Tsumoto and P. Novais, Argumentation with goals for clinical decision support in multimorbidity, in: *17th International Conference on Autonomous Agents and Multiagent Systems*, IFAAMAS, Stockholm, 2018, pp. 2031–2033.

[74] S. Parsons, C. Sierra and N. Jennings, Agents that reason and negotiate by arguing, *Journal of Logic and Computation* **8**(3) (1998), 261–292. ISSN 0955-792X. doi:10.1093/logcom/8.3.261.

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

[76] H. Prakken, An abstract framework for argumentation with structured arguments, *Argument & Computation* **1**(2) (2010), 93–124. doi:10.1080/19462160903564592.

[77] H. Prakken, A top-level model of case-based argumentation for explanation, in: *2nd International Workshop on Dialogue, Explanation and Argumentation for Human–Agent Interaction*, 2020.

[78] M.A. Qassas, D. Fogli, M. Giacomin and G. Guida, ArgMed: A support system for medical decision making based on the analysis of clinical discussions, in: *Real-World Decision Support Systems: Case Studies*, J. Papathanasiou, N. Ploskas and I. Linden, eds, Springer, 2016, pp. 15–41. doi:10.1007/978-3-319-43916-7_2.

[79] A. Rago, O. Cocarascu and F. Toni, Argumentation-based recommendations: Fantastic explanations and how to find them, in: *27th International Joint Conference on Artificial Intelligence*, International Joint Conferences on Artificial Intelligence Organization, Stockholm, 2018, pp. 1949–1955. doi:10.24963/ijcai.2018/269.

[80] I. Rahwan and G.R. Simari, *Argumentation in Artificial Intelligence*, Springer, 2009, pp. 1–487. ISBN 978-0-387-98196-3. doi:10.1007/978-0-387-98197-0.

[81] A. Raymond, H. Gunes and A. Prorok, Culture-based explainable human-agent deconfliction, in: *19th International Conference on Autonomous Agents and MultiAgent Systems*, A.E.F. Seghrouchni, G. Sukthankar, B. An and N. Yorke-Smith, eds, IFAAMAS, Auckland, 2020, pp. 1107–1115.

[82] D. Riaño and W. Ortega, Computer technologies to integrate medical treatments to manage multimorbidity, *Journal of Biomedical Informatics* **75** (2017), 1–13. 1532-0464 (Electronic) 1532-0480 (Linking). doi:10.1016/j.jbi.2017.09.009.

[83] L. Sacchi, S. Rubrichi, C. Rognoni, S. Panzarasa, E. Parimbelli, A. Mazzanti, C. Napolitano, S.G. Priori and S. Quaglini, From decision to shared-decision: Introducing patients' preferences into clinical decision analysis, *Artificial Intelligence in Medicine* **65**(1) (2015), 19–28. ISSN 0933-3657. doi:10.1016/j.artmed.2014.10.004.

[84] C. Sakama, Abduction in argumentation frameworks, *Journal of Applied Non-Classical Logics* **28**(2–3) (2018), 218–239. doi:10.1080/11663081.2018.1487241.

[85] I. Sassoon, N. Kökciyan, E. Sklar and S. Parsons, Explainable argumentation for wellness consultation, in: *1st International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, Vol. **11763**, D. Calvaresi, A. Najjar, M. Schumacher and K. Främling, eds, Springer, Montreal, 2019, pp. 186–202. doi:10.1007/978-3-030-30391-4_11.

[86] C. Schulz and F. Toni, ABA-based answer set justification, *Theory and Practice of Logic Programming* **13**(Online – Supplement) (2013), 4–5.

[87] N. Sendi, N. Abchiche-Mimouni and F. Zehraoui, A new transparent ensemble method based on deep learning, *Procedia Computer Science* **159** (2019), 271–280. doi:10.1016/J.PROCS.2019.09.182.

[88] D. Šešelja and C. Straßer, Abstract argumentation and explanation applied to scientific debates, *Synthese* **190**(12) (2013), 2195–2217. doi:10.1007/s11229-011-9964-y.

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

[90] E.I. Sklar and M.Q. Azhar, Explanation through argumentation, in: *6th International Conference on Human–Agent Interaction*, M. Imai, T. Norman, E. Sklar and T. Komatsu, eds, ACM, Southampton, 2018, pp. 277–285. ISBN 9781450359535. doi:10.1145/3284432.3284470.

[91] M. Spiotta, P. Terenziani and D.T. Dupré, 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. doi:10.1109/TKDE.2017.2734084.

[92] P.M. Thang and H.T. Luong, Translating preferred subtheories into structured argumentation, *Journal of Logic and Computation* **24**(4) (2014), 831–850. doi:10.1093/logcom/ext049.

[93] S.T. Timmer, J.-J.C. Meyer, H. Prakken, S. Renooij and B. Verheij, A two-phase method for extracting explanatory arguments from Bayesian networks, *International Journal of Approximate Reasoning* **80** (2017), 475–494. doi:10.1016/j.ijar.2016.09.002.

[94] P. Tolchinsky, U. Cortés, S. Modgil, F. Caballero and A. López-Navidad, Increasing human-organ transplant availability: Argumentation-based agent deliberation, *IEEE Intelligent Systems* **21**(6) (2006), 30–37. doi:10.1109/MIS.2006.116.

[95] F. Toni, A tutorial on assumption-based argumentation, *Argument & Computation* **5**(1) (2014), 89–117. doi:10.1080/19462166.2013.869878.

[96] R. Tsopra, J.B. Lamy and K. Sedki, Using preference learning for detecting inconsistencies in clinical practice guidelines: Methods and application to antibiotherapy, *Artificial Intelligence in Medicine* **89** (2018), 24–33. doi:10.1016/j.artmed.2018.04.013.

[97] B. Verheij, DefLog: On the logical interpretation of prima facie justified assumptions, *Journal of Logic and Computation* **13**(3) (2003), 319–346. doi:10.1093/logcom/13.3.319.

[98] B. Verheij, Artificial argument assistants for defeasible argumentation, *Artificial Intelligence* **150**(1–2) (2003), 291–324, https://linkinghub.elsevier.com/retrieve/pii/S0004370203001073. doi:10.1016/S0004-3702(03)00107-3.

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

[100] T. Wakaki, Assumption-based argumentation equipped with preferences, in: *Principles and Practice of Multi-Agent Systems – 17th International Conference*, H.K. Dam, J.V. Pitt, Y. Xu, G. Governatori and T. Ito, eds, Lecture Notes in Computer Science, Vol. 8861, Springer, Gold Coast, 2014, pp. 116–132. doi:10.1007/978-3-319-13191-7_10.

[101] T. Wakaki, K. Nitta and H. Sawamura, Computing abductive argumentation in answer set programming, in: *6th International Workshop on Argumentation in Multi-Agent Systems*, P. McBurney, I. Rahwan, S. Parsons and N. Maudet, eds, Springer, Budapest, 2009, pp. 195–215. doi:10.1007/978-3-642-12805-9_12.

[102] D. Walton, *Argumentation Schemes for Presumptive Reasoning*, L. Erlbaum Associates, 1996. ISBN 080582071X.

[103] D. Walton, A new dialectical theory of explanation, *Philosophical Explorations* **7**(1) (2004), 71–89. doi:10.1080/1386979032000186863.

[104] D. Walton, Explanations and arguments based on practical reasoning, in: *Explanation-Aware Computing, 2009 IJCAI Workshop*, Pasadena, T. Roth-Berghofer, N. Tintarev and D.B. Leake, eds, 2009, pp. 72–83.

[105] S. Wilk, M. Michalowski, W. Michalowski, D. Rosu, M. Carrier and M. Kezadri-Hamiaz, Comprehensive mitigation framework for concurrent application of multiple clinical practice guidelines, *Journal of Biomedical Informatics* **66** (2017), 52–71. ISSN 1532-0464. doi:10.1016/j.jbi.2016.12.002.

[106] L. Xiao and J. Fox, A distributed decision support architecture for the diagnosis and treatment of breast cancer, in: *International Conference on Health Information Science*, G. Huang, X. Liu, J. He, F. Klawonn and G. Yao, eds, Lecture Notes in Computer Science, Vol. 7798, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016, pp. 9–21. ISBN 978-3-642-37898-0. doi:10.1007/978-3-319-48335-1_2.

[107] A.P. Young, N. Kökciyan, I. Sassoon, S. Modgil and S. Parsons, Instantiating metalevel argumentation frameworks, in: *7th International Conference on Computational Models of Argument*, S. Modgil, K. Budzynska and J. Lawrence, eds, IOS Press, Warsaw, 2018, pp. 97–108. doi:10.3233/978-1-61499-906-5-97.

[108] V. Zamborlini, M. da Silveira, C. Pruski, A. ten Teije, E. Geleijn, M. van der Leeden, M. Stuiver and F. van Harmelen, Analyzing interactions on combining multiple clinical guidelines, *Artificial Intelligence in Medicine* **81** (2017), 78–93. ISSN 1873-2860 (Electronic) 0933-3657 (Linking). doi:10.1016/j.artmed.2017.03.012.

[109] V. Zamborlini, R. Hoekstra, M. Da Silveira, C. Pruski, A. ten Teije and F. van Harmelen, Inferring recommendation interactions in clinical guidelines, *Semantic Web* **7**(4) (2016), 421–446. doi:10.3233/SW-150212.

[110] Z. Zeng, X. Fan, C. Miao, C. Leung, C. Jing Jih and O. Yew Soon, Context-based and explainable decision making with argumentation, in: *17th International Conference on Autonomous Agents and MultiAgent Systems*, IFAAMAS, Stockholm, 2018, pp. 1114–1122.

[111] Z. Zeng, C. Miao, C. Leung and C.J. Jih, Building more explainable artificial intelligence with argumentation, in: *32nd AAAI Conference on Artificial Intelligence (Doctoral Consortium)*, S.A. McIlraith and K.Q. Weinberger, eds, AAAI Press, New Orleans, Louisiana, 2018.

[112] Q. Zhong, X. Fan, X. Luo and F. Toni, An explainable multi-attribute decision model based on argumentation, *Expert Systems with Applications* **117** (2019), 42–61. doi:10.1016/j.eswa.2018.09.038.