

# Argumentation-based Explanations of Multimorbidity Treatment Plans

Qurat-ul-ain Shaheen<sup>1</sup>[0000-0003-2258-4476], Alice Toniolo<sup>2</sup>[0000-0002-6816-63604], and  
Juliana K. F. Bowles<sup>2</sup>[0000-0002-5918-9114]

<sup>1</sup> IIIA-CSIC, Barcelona, 08193, Spain  
qurat@iia.csic.es

<sup>2</sup> Computer Science, University of St Andrews, St Andrews KY16 9SX, UK  
{a.toniolo, jkfb}@st-andrews.ac.uk

**Abstract.** We present an argumentation model to explain the optimal treatment plans recommended by a Satisfiability Modulo Theories solver for multimorbid patients. The resulting framework can be queried to obtain supporting reasons for nodes on a path following a model of argumentation schemes. The modelling approach is generic and can be used for justifying similar sequences.

**Keywords:** Multimorbidity · Argumentation schemes · SMT solvers

## 1 Introduction

Multimorbidity, the presence of two or more chronic conditions in a person, is a major healthcare concern across the globe. In the UK, NICE<sup>3</sup> publishes evidence-based recommendations for treatment, called *Clinical Pathways*, to formalise treatment practices. However, these pathways are based on evidence from single health conditions and rarely consider the presence of comorbidities. Hence, treating multimorbidity involves managing adverse interactions between several drugs corresponding to different conditions. Several efficient approaches have been developed based on Satisfiability Modulo Theories (SMT) solvers [2] to identify safe treatment plans. Medical practitioners are, however, naturally reluctant to use automatic recommendations if they do not understand the underlying reasoning. In this work, we combine the optimising efficiency of SMT solvers with an argumentation framework for explaining their recommendations. While the role of argumentation has been investigated in healthcare [4], explanations in healthcare [8] and multimorbidity [3], the contribution of this paper is a novel approach for modelling graphs and their explanations using argumentation in the context of multimorbidity. The SMT solver provides an optimal solution indicating the safest medical treatment plan for multimorbidity. A set of argumentation schemes captures reasons for these medical treatments as well as possible inconsistencies between recommended plans. We then show how critical questions can be used to query an argumentation framework, where an agent can provide interactive explanations to a medical practitioner. In this paper, we present the argumentation framework and its underlying schemes which can be used in a multi-agent system to uncover explanations of recommended paths.

<sup>3</sup> For details see <https://www.nice.org.uk>

## 2 Argumentation and Multimorbidity for NICE Pathways

Clinical pathways are generic flowchart representations. They include guidelines for drug treatments of specific conditions as well as other best practices. We refer to them as *NICE pathways* in this work. Kovalov and Bowles [5] use the SMT solver Z3 [7] to find the safest path across multiple NICE pathways such that adverse drug interactions between all the NICE pathways under consideration are minimised. SMT solvers solve Boolean Satisfiability problems by finding a truth assignment modulo theories and its optimal assignment for some defined objective function. Three types of adverse interactions are considered in [5]: between drugs, between a drug and a disease, and patient intolerance for specific drugs. The approach is scalable to any number of drug alternatives and finds the optimal solution minimising conflicts according to given criteria as opposed to an all or nothing approach. While the solver captures the NICE pathways quantitatively to indicate progression to the next stage of treatment, as a black-box model it does not provide the underlying qualitative intuition. In this paper, our focus is the encoding of this qualitative information in the solver’s output.

We adopt the formal representation of [2, 5] to represent Pharmaceutical Graphs (PG) which capture the original ordering and branching structure of the NICE pathway in terms of the recommended drug or drug groups. A PG is a directed acyclic graph with a root node for a disease, and where every other node in the graph is associated with a drug or a drug group. For patients with multimorbidities, several event structures are used to model the NICE pathways of each of the conditions they have. A path in the graph represents a complete treatment plan for a single disease, and a combination of paths across diseases is extracted in [5] via an objective function on the basis of a total score for Z3. This score is a combination of medicine efficacy (positive score) and drug interaction conflict (negative score) permitting to identify the optimal treatment plan.

Our definition of PG is based on that of [2], formally considering the unfolding of a PG as an event structure where events are associated with a drug or a group and different treatment options are captured by conflicting events.

**Definition 1.** *An event structure is a triple  $E = (Ev, \rightarrow^*, \#)$  where  $Ev$  is a set of events and  $\rightarrow^*, \# \in Ev \times Ev$  are binary relations called causality and conflict, respectively. Causality  $\rightarrow^*$  is a partial order. Conflict  $\#$  is symmetric and irreflexive, and propagates over causality, i.e.,  $(e\#e' \wedge e' \rightarrow^* e'') \Rightarrow e\#e''$  for all  $e, e', e'' \in Ev$ . Two events are concurrent, written  $e \parallel e'$ , iff  $\neg(e \rightarrow^* e' \vee e' \rightarrow^* e \vee e\#e')$ .*

Here we consider PG as the graph formed by the causality relationships of  $E$  such that  $PG = (Ev, \rightarrow^*)$ , and our starting representation of a PG is correspondent to the representation of Z3 input as event structure. We refer to a trace of execution on PG as a sequence of events  $Ppath = \langle e_0; \dots; e_n \rangle$ ,  $Ppath \subseteq PG$ , where for each  $e_i, e_j \in Ev$  there exists a causal relation  $(e_i \rightarrow^* e_j) \in \rightarrow^*$ . Hence the output from Z3 is a set of paths for each disease optimising a goal over all sets of event structures under consideration.

In order to compute justifications for this path, we define an argumentation model using a simpler version of ASPIC+ [6] without preferences, strict rules and undercuts. An argumentation framework  $AF$  is formed by arguments  $\mathcal{A}$  and attacks  $\rightarrow: \mathcal{A} \times \mathcal{A}$ . An argument  $A \in \mathcal{A}$  is constructed from a knowledge base  $K \subseteq \mathcal{L}$  where  $\mathcal{L}$  is a propositional logic language, and defeasible rules  $\mathcal{R}$ .  $A \in \mathcal{A}$  is a premise in  $K$  or a derivation

of a conclusion from rules and premises in  $K$ . Argument  $A1$  attacks  $A2$ ,  $A1 \rightarrow A2$ , if the conclusion of  $A1$  is contrary to a premise (undermines) or to a conclusion (rebuts) of  $A2$  determined by a contrariness function  $\bar{\cdot}: \mathcal{L} \rightarrow 2^{\mathcal{L}}$ . The sets of acceptable arguments (i.e., extensions  $\xi$ ) in an  $AF$  is computed according to the preferred semantics. A model of argumentation schemes, stereotypical patterns of reasoning, defines premises, conclusions and critical questions [10]. Schemes are used to derive grounded defeasible rules  $k_i \in \mathcal{R}$  [6]. In our work, attacks will be derived through ASPIC+, while critical questions are used to explore arguments within an extension.  $\mathcal{L}$  contains the set of propositions  $Props$  in turn divided into subsets  $\mathcal{T}_i \subset Props$ . A lower case letter refers to a proposition type  $t_i \in \mathcal{T}_i$  and we refer to grounded propositions with upper case letters. An argument scheme is defined as  $Arg_\gamma: t_0, \dots, t_j \Rightarrow t_n$  where  $\gamma$  is the name of the scheme and  $t_i \in \mathcal{T}_i$  from which we can derive a rule  $k_\gamma: P_0, \dots, P_j \Rightarrow P_n, k_\gamma \in \mathcal{R}$ .

### 3 Argumentation Model

In our model, we adapt PG, input to Z3, for adding explanations and convert it into its equivalent argument representation. Next, we apply a knowledge engineering process to extract the explanations from NICE pathways that were left out by the SMT solver. Lastly, we link the explanations to the respective PG nodes in our final model.

**Convert Pharmaceutical Graph to PGraph.** Starting from a PG, the event structure representation presented in Section 2, we define a *PGraph* as a graph extension of PG. *PGraph* differentiates from its parent PG by having the graph structure marked explicitly in terms of branches required for explanation modelling. A *PGraph*  $= (Ev, \rightarrow^*)$  augments the set of events with a new set of nodes where *Sb* or *StartBranch* is a set of events marking the start of a branch, *Eb* or *EndBranch* marks the end of a branch. To represent drug groups as sub-branches in PG, we consider a set of nodes *Dg* or *Drug-Group* named after the drug group which marks the start of the sub-branch. Similarly we have a set of nodes *Eg* or *EndGroup* to mark the end of the sub-branch. Causality relationships link the new nodes to their respective branches. A path on a PG is then mapped to a path on *PGraph* by accounting for additional marking nodes. In Fig. 1 we show an example of  $PG = (Ev, \rightarrow^*)$  for Diabetes from [5] and its correspondent extension *PGraph*. The root node represents the disease (Diabetes), while the remaining nodes represent recommended drugs or drug groups. For example, the node Sulfonylurea indicates that seven drugs belong to this group.

**Modelling Paths in the Argumentation Model.** A casual relation  $e \rightarrow^* e'$  in *PGraph* is represented using a Transition Argument Scheme. Transition arguments accumulate the path from start to the current transition using a linear argument structure.

**Definition 2.** A Transition Argument Scheme is  $Arg_T = \{ \text{Given } e \text{ in a } PGraph, e' \text{ follows from } e \}$  where  $e \rightarrow^* e' \in PGraph$ .

This scheme can be represented as  $Arg_T: e \Rightarrow e'$ , where  $e, e' \in \mathcal{T}_{Ev}$  are propositions of type event and  $\mathcal{T}_{Ev}$  contains all events  $Ev$  of *PGraph*. An inference rule derived through  $Arg_T$  can be written as  $k_{T_1}: E1 \Rightarrow E2$  with  $k_{T_1} \in \mathcal{R}$ . A single critical question

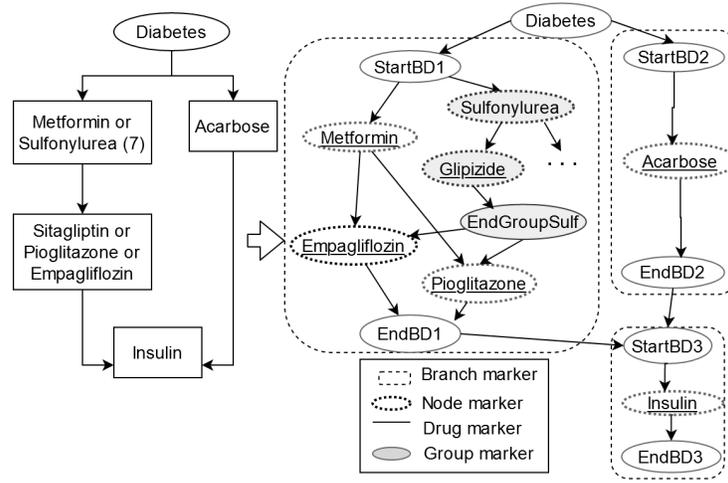


Fig. 1. PG to PGraph conversion for Diabetes.

is defined:  $\mathbf{CQ}_T$  – Is there a transition to  $e$ ? The question undermines the argument on its premise with  $\neg e$ , and other arguments of type  $Arg_T$  can be used to respond.

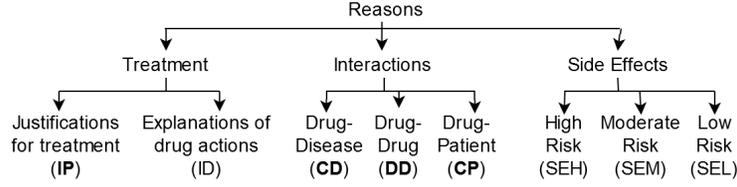
In order to model different paths taken along PGraph as alternative solutions, here we introduce a negation to represent alternatives such that for all concurrent events  $e', e'' \in Ev$ ,  $e' \parallel e''$  we have  $e' \in \bar{e}'', e'' \in \bar{e}'$ .

**Structured Explanation for Clinical Pathways.** We base the explanations of the recommended paths on the drug information provided by NICE. We assess the practical value of this content by talks with experts. Then we manually filter out relevant information and organise it into different levels and types.

*Level:* The choice of a node in a graph involves decisions at different abstraction levels. For example, starting from the initial node, the medical practitioner needs to choose a branch. Next, they decide which node to transition to and why. This is important because in the treatment plans, moving to the next node in the graph indicates that the previous treatment is no longer effective, so it is discouraged. Once made, the transition as well as the choice of drug or drug group needs to be justified. Hence, the abstraction levels represent different types of decisions to be made. Higher levels of abstraction act as filters for lower levels. As the choice of a node moves from higher to lower level, explanations become more specific and adapt to the user’s need for detail. For example, a non-expert might be satisfied with a high-level explanation but a more expert user might press for more specific reasons encoded in the lower levels.

We identify four levels of abstraction for PGraph (cf. Fig. 1): a *Branch* contains *Nodes* where a *Node* can be a *Drug* or a composite *Group*, composed of *Drug* nodes.

1. *Branch Level:* Given a PGraph, the first decision is to choose between alternative branches. This level justifies a transition to a specific branch.



Reasons $\mathcal{T}_{Reas}$	Maps to	$\mathcal{T}_{Type}$	$\mathcal{T}_{Level}$
R1	it is indicated for age less than 55 years	<b>IP</b>	<b>BRANCH</b>
R2	it is the default treatment on the recommended pathway	<b>IP</b>	<b>BRANCH</b>
R3	it is indicated in next stage of treatment if review of previous treatment does not work	<b>IP</b>	<b>NODE</b>
R4	it is contraindicated in	<b>CD</b>	<b>DRUG</b>
R5	it adversely reacts with suggested drug	<b>DD</b>	<b>DRUG</b>

**Fig. 2.** Type hierarchy for Treatment, Interactions and Side Effects with examples for Diabetes.

2. *Node Level*: This level allows reasoning about why a transition is taken irrespective of the destination node. In the context of multimorbidity, these reasons allow talking about transitions into the next stage of treatment.
3. *Drug Level*: Reasons about the destination node, i.e. a drug or a drug group.
4. *Group Level*: Reasons about relationships between drugs in a drug family.

In order to identify what level of explanation a reason is attributed to, we define specific propositions in our language  $\mathcal{T}_{Level} = \{BRANCH, NODE, DRUG, GROUP\} \subset Props$ .

*Type*: We classify the reasons for choosing or not choosing a drug into four types based on the content. These are distributed across different levels: 1) Clinical pathway; 2) Treatment; 3) Interactions; and 4) Side Effects. The Clinical Pathway type allows the *PGraph* structure itself to be used as justification for a drug choice. Figure 2 shows the sub-classification within the remaining types with those used in this work highlighted in bold. As above, a set of dedicated propositions are used to identify the different type of reason:  $\mathcal{T}_{Type} = \{IP, CD, CP, DD, PGraph\} \subset Props$ .

*Reason*: represents the explanation text and is associated with a *Level* and a *Type*. A reason is identified with a proposition type  $r_i \in \mathcal{T}_{Reas} \subset Props$ . The table in Fig. 2 provides some examples of reasons indicating the proposition type, level and a description.

**Argument Schemes for Explanations.** The *Explain Argument Scheme* models explanations based on the reasons, types and levels and shows a support or counter reason for giving a drug represented as event  $e$ .

**Definition 3.** An *Explain Argument Scheme*  $Arg_E$ , is such that given a reason  $r \in \mathcal{T}_{Reas}$  with attributes  $t \in \mathcal{T}_{Type}$  and  $l \in \mathcal{T}_{Level}$  for event  $e$  in *PGraph*, it follows that there is a justification for prescribing  $e$ .

We customize the generic  $Arg_E$  to model different types of supporting explanations that can highlight adverse interactions between events in the *PGraph*. We specify additional elements as follows in order to define these schemes. We assume  $\mathcal{T}_{Dis} = \{d|d \text{ is a disease name}\}$ ,  $\mathcal{T}_{Cond} = \{c|c \text{ is a medical condition expressing patient state}\}$ ,  $\mathcal{T}_{Exp} = \{explain_e|explain_e \text{ is a proposition representing an explanation for a drug/drug}$

group  $e \in \mathcal{T}_{Ev}$  and  $\mathcal{T}_{Dis}, \mathcal{T}_{Cond}, \mathcal{T}_{Exp} \subset Props$ . The general argument scheme is presented as  $Arg_E : r, t, l \Rightarrow explain_e$ . Four types of schemes  $Arg_E$  are defined on the basis of the type of reason:  $Arg_{EIP}$  mentions a justification for a treatment;  $Arg_{EDD}$  indicates that a drug  $e_1$  is prescribed as another drug  $e_2$  reacting with  $e_1$  has not been prescribed;  $Arg_{ECD}$  indicates that  $e$  cannot be taken in case of a disease  $s \in \mathcal{T}_{Dis}$  but the disease is not present; and  $Arg_{ECP}$  includes information on whether  $e$  can or cannot be taken in presence or absence of another condition  $c \in \mathcal{T}_{Cond}$ .

Reasons from  $Arg_E$  do not aggregate. A drug can be prescribed as long as there is a single surviving  $Arg_E$  to justify it. *Explain* arguments allow a critical question such as: **CQ<sub>E</sub>** – *Is there an alternative combination of  $r, t, l$  for there being a justification for  $e$ ?* This rebuts the conclusion with  $\neg explain_e$ , and can be answered with another  $Arg_E$ .

**Linking the Two Models.** The *Give Argument Scheme* links graph and explanations:

**Definition 4.**  $Arg_G = \{Given\ an\ event\ e\ in\ PGraph\ and\ an\ explanation\ explain_e \in \mathcal{T}_{Exp}\ in\ favour\ of\ e,\ it\ follows\ that\ e\ can\ be\ prescribed\}$ .

In order to define formally this scheme, we use  $\mathcal{T}_{Pres} = \{give_e | give_e\ is\ a\ proposition\ expressing\ prescription\ of\ e \in \mathcal{T}_{Ev}\}$ . A *Give* argument scheme is then formulated as  $Arg_G : e, explain_e \Rightarrow give_e$  where  $give_e \in \mathcal{T}_{Pres}$ . Two critical questions are defined for this scheme: **CQ<sub>G1</sub>** – *Is there a transition to  $e$ ?*; and **CQ<sub>G2</sub>** – *Is there an explanation for  $e$ ?* Both CQs challenge the premises:  $CQ_{G1}$  (same as  $CQ_T$ ) with  $\neg e$  and  $CQ_{G2}$  with  $\neg explain_e$  which can be responded with  $Arg_T$  and  $Arg_E$ , respectively.

**General Argumentation Model.** So far, we have outlined a number of argumentation schemes to derive defeasible rules  $\mathcal{R}$ . The knowledge base is defined as  $K \subseteq \mathcal{T}_{Dis} \cup \mathcal{T}_{Cond} \cup \mathcal{T}_{Level} \cup \mathcal{T}_{Type} \cup \mathcal{T}_{Reas} \cup \mathcal{T}_{Ev}$  where  $\mathcal{T}_{Ev}$  represents a subset of negated drug/drug group events, s.t.  $\neg e \in \mathcal{T}_{Ev}, e \in \mathcal{T}_{Ev}$ . The components of  $K$  are identified as: (i) for  $q_i \in (\mathcal{T}_{Dis} \cup \mathcal{T}_{Cond} \cup \mathcal{T}_{Ev})$ , we assert  $q_i \in K$  based on NICE pathways, reflecting the domain properties, which will constitute premises for *Explain* Arguments; (ii) for  $q_i \in (\mathcal{T}_{Level} \cup \mathcal{T}_{Type} \cup \mathcal{T}_{Reas})$ , we assert  $q_i \in K$  as part of the modelling permitting an administrator to tune the explanations returned by the system.  $\mathcal{R}$  and  $K$  form the basis for an *AF* to explain an event structure representing a NICE treatment pathway.

Asserting the disease node in the graph activates the graph through the linear structure of *Transition* arguments. At this point we can derive a set of preferred extensions  $\xi$  from the argumentation system as described in Section 2. A preferred extension is then  $\xi_i = \{A_{T_1}, \dots, A_{T_n}, A_{E_1}, \dots, A_{E_k}, A_{G_1}, \dots, A_{G_m}\}$  where  $A_{T_i} \in \mathcal{A}$  are arguments about the treatment path,  $A_{E_i} \in \mathcal{A}$  are justifications for prescribing the treatment,  $A_{G_i} \in \mathcal{A}$  are treatments that can be justifiably prescribed. Hence,  $\xi_i$  maps an SMT solver output  $Ppath = \langle e_0; \dots; e_n \rangle$  and additionally incorporates corresponding explanations.

*Transition* arguments rebut other *Transition* arguments and undermine corresponding *Give* and *Explain* arguments on their *Transition* premises. Consequently, *Give* arguments can be undermined on their *Transition* premise directly and indirectly by undermining one of the premises of their *Explain* arguments.

**Example of Argumentation Model.** Here we show part of the ASPIC+ model for the Diabetes *PGraph* from Fig. 1. The adverse interactions modelled in this example are only for demonstrative purposes. The example can be extended to include multiple disease *PGraphs* and their explanations by adding corresponding premises, rules and contrariness definitions. The resulting extensions will then include combinations of paths across all graphs, representing the safest path in the context of multimorbidity.

We now show how the *AF* can be used by an agent to justify prescribed drugs. Assuming the SMT solver gives the optimal path as branch 1 of Diabetes *PGraph* in Fig. 1. We map this to  $Ppath = \langle Diabetes; StartBD1; Metformin; Pioglitazone; EndBD1 \rangle$  and get the corresponding preferred extension  $\xi_1 = \{A60, A54, A56, A57, ..\}$ . Figure 3 shows the attacks in the resulting argument graph for this extension along with a subset of the relevant argument mappings. The arguments included in the extension are highlighted in grey. Assume the agent proposes to give Pioglitazone, *A60*. A user can ask for justifications using some of the critical questions. The agent responds with an argument that can resolve the challenge. For example, if the user asks for a justification about the transition to Pioglitazone,  $CQ_{G1}$ , the agent can respond with *Because there is a transition to Pioglitazone from Metformin*, via argument *A54*. If the user asks about an explanation for this drug,  $CQ_{G2}$ , the agent can respond with *A56: Because it is indicated for age less than 55 years (R1)*. The agent can move *A57, Because it is indicated in the next stage of treatment if a review of previous treatment does not work (R3)*, in response to a request for alternative explanations to the drug,  $CQ_{E1}$ .

## 4 Conclusion and Future Work

We presented an argumentation model to justify the choices on a clinical pathway, as identified by the SMT solver for multimorbid patients. While the solver only encoded drug information in the optimal path, our system delivers additional information extracted from the NICE pathways as well as other sources that justifies each event in the path. Our approach can generalise to explanations of similar graph structures.

A recent approach has integrated argumentation with the Transition-based Medical Recommendation (TMR) model to represent guidelines and their interactions [3]. While they use argumentation for reasoning, we present a novel approach to show how argumentation can add interpretability to a black-box model such as a SMT solver. Although significant work has been invested in making machine learning models more transparent [1], the same focus on making SMT solvers transparent has lacked so far.

The explanations from our model can be used as the basis of an agent dialogue protocol which provides justifications to the user in an interactive way [9], similar to the approach of [8]. Possible future directions include conducting usability studies to see how useful are the generated explanations and formulating the content of the explanations on data collected from patients and medical practitioners.

## References

1. Biran, O., Cotton, C.: Explanation and justification in machine learning: A survey. In: Proceedings of the IJCAI-17 Workshop on Explainable AI (XAI). vol. 8, pp. 8–13 (2017)

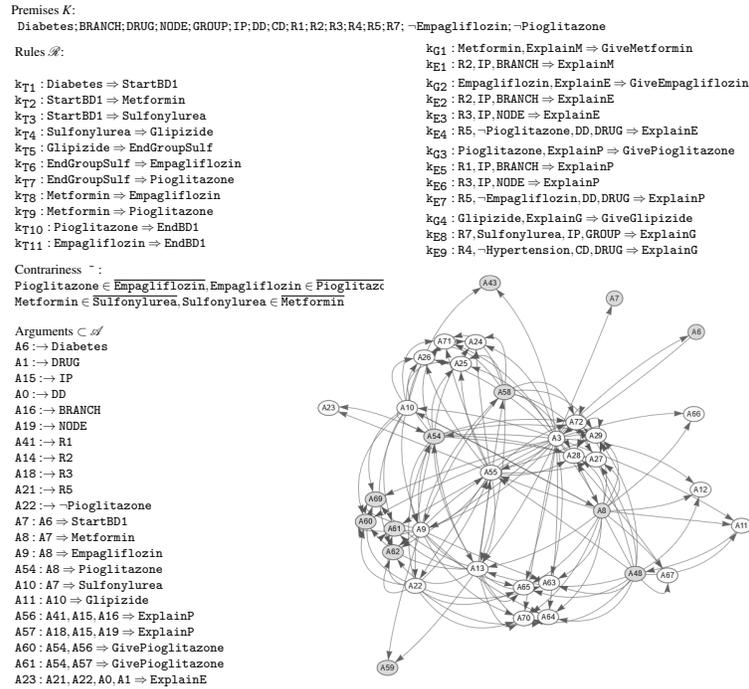


Fig. 3.  $AF$  for branch 1 of Diabetes  $PGraph$  from Figure 1.

- Bowles, J., Caminati, M., Cha, S., Mendoza, J.: A framework for automated conflict detection and resolution in medical guidelines. *Science of Computer Programming* **182**, 42 – 63 (2019). <https://doi.org/https://doi.org/10.1016/j.scico.2019.07.002>
- Čyras, K., Oliveira, T.: Resolving Conflicts in Clinical Guidelines using Argumentation. In: *Proc. of Autonomous Agents and Multiagent Systems*. pp. 1731–1739 (2019)
- Hunter, A., Williams, M.: Aggregating evidence about the positive and negative effects of treatments. *Artificial Intelligence in Medicine* **56**(3), 173–190 (2012)
- Kovalov, A., Bowles, J.K.F.: Avoiding medication conflicts for patients with multimorbidities. In: *Integrated Formal Methods*. pp. 376–390. Springer (2016)
- Modgil, S., Prakken, H.: The ASPIC+ framework for structured argumentation: A tutorial. *Argument & Computation* **5**(1), 31–62 (2014)
- de Moura, L., Bjørner, N.: Z3: An efficient SMT solver. In: Ramakrishnan, C.R., Rehof, J. (eds.) *Tools and Algorithms for the Construction and Analysis of Systems*. pp. 337–340. Springer B.H. (2008)
- Sassoon, I., Kökciyan, N., Sklar, E., Parsons, S.: Explainable argumentation for wellness consultation. In: *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*. pp. 186–202. Springer (2019)
- Shaheen, Q.u.a., Toniolo, A., Bowles, J.K.F.: Dialogue games for explaining medication choices. In: *Rules and Reasoning (RuleML+RR 2020)*. LNCS 12173, Springer (2020)
- Walton, D., Reed, C., Macagno, F.: *Argumentation schemes*. Cambridge Univ. Press (2008)