HHAI 2023: Augmenting Human Intellect
P. Lukowicz et al. (Eds.)

doi:10.3233/FAIA230125

Using Non-Monotonic Logics to Create a Dynamic Framework for a Behavior Support Agent

Johanna WOLFF
j.d.wolff@utwente.nl

Abstract. Behavior support agents can assist humans in accomplishing a variety of goals by suggesting actions that promote the desired outcome while being in line with the user's needs and preferences. In order to make these agents more effective, flexible and responsible, this research aims to create a framework which allows for more interaction between the agent and the user. By using techniques from non-monotonic reasoning, this work aims to model the knowledge base of the agent so that it aligns with the user's mental model and is able to be modified by the user through new input. In order for the agent to be able to explain its output to the user, the reasoning process needs to be explicit and traceable, which this work intends to incorporate into a logical framework.

Keywords. Non-Monotonic Logics, Knowledge Representation, Human-Machine Alignment, User Modeling, Behavior Support Agents

1. Introduction

Technology that is intended to support users in achieving their goals by suggesting certain behavior has already become ingrained in many people's lives [1]. Examples for this include agents that help the user schedule meetings, encourage them to exercise and eat healthy or assist them in saving energy within a smart home environment. In order for these behavior support agents to effectively support the user, especially over a longer period of time, they need to be able to adapt to their user's goals, capabilities and preferences [2]. The use of machine learning techniques has made it possible to create artificial agents that can optimize their functionality based on the previous behavior of the user, leading to a high degree of personalization. However, these data-driven approaches make it difficult for the user to directly access or even influence the knowledge base of the agent since relevant concepts are not explicitly represented. Therefore, if the user wants to change their behavior or priorities, it is often unclear how to input these changes into the agent and what effects they might have on the agent's output.

Additionally, there is a growing desire to ensure that artificial agents are designed responsibly and the user remains in control of how they use the technology [3]. Within the field of Hybrid Intelligence [4], these issues are addressed by emphasizing the importance of interaction and collaboration between humans and

AI agents. In the context of behavior support agents the user should therefore be able to continuously update the agent's knowledge base in order to be in control of the information that the agent reasons with. This means that the user can directly change, add or delete information within the agent's internal model.

On the other hand, the agent should be able to explain the reasoning process that leads to conclusions or potential conflicts. The importance of this is not only mentioned in voluntary AI design guidelines like [5], but also in laws such as Regulation (EU) 2016/679 (EU, 2016) better known as GDPR. An explainable agent can help the user understand and trust its suggestions [6], [7], [8]. Moreover, the agent should be capable of asking the user for additional input when it recognizes a conflict or gap in its knowledge base.

We will study the underlying structures of these interactions between agent and user by looking at the types of information that occur and how they can be incorporated in and extracted from the knowledge base of the agent. The use of knowledge-based methods allows us to represent both the knowledge base and the reasoning process explicitly, which potentially makes them easier to be understood and influenced by the user. We will specifically investigate how techniques from non-monotonic reasoning can be combined with methods from knowledge representation and reasoning in order to create a logical framework for an agent that supports the interactions mentioned above. Non-monotonic logics have been studied a lot within the field of artificial intelligence because they can formalize aspects of human common-sense reasoning [9]. In our work, we plan to use these reasoning techniques to model an agent in a way that aligns with the user's mental model of the situation. We will give a brief overview of the most common methods from non-monotonic reasoning in Section 3.

An explicit example of a possible application of our research is a diabetes support agent that is intended to assist a user who has either been diagnosed with diabetes, or is at risk of developing diabetes. In both cases the agent is intended to help the user change their lifestyle in order to stay healthy. Possible actions of the agent may include suggesting an exercise schedule or other ways to remain active, keeping track of the user's diet and monitoring medical data. Being required to change their entire lifestyle immediately can often feel overwhelming, so the user may benefit from an agent which is able to be flexible in its goals and can adapt them over time. Additionally, such a lifestyle change needs to be long-lasting. The user's priorities and preferences may change over time, for example if starting a new job requires them to change their schedule or they choose to prioritize time with their family over working out. While the agent should be able to accommodate these changes, it is also important that certain safety features are always guaranteed, such as the interpretation of medical data or the priority of the user's life being higher than any goals the user may add. Since diabetes is a disease which affects a large number of people, many of whom have low health and tech literacy, it is important for such an agent to be able to explain its suggestions in a way that encourages the user to trust and follow them. Potential conflicts between the user's input and the agent's knowledge need to be communicated in order to discourage the user from potentially dangerous behavior, for example if the user expresses goals which would negatively impact their health.

2. Problem Statement

Our goal is to develop a logic which can be used for a personal support agent that is able to adapt to the needs of a user and explain its conclusions through interaction with the user. The agent will use the logic to reason about the information it has available in order to determine how best to support the user. This may include suggesting possible actions to the user, giving the user new information or asking the user for additional input. The user should be able to give additional input whenever they want to add or change the information that the agent uses. The agent should also be able to communicate with the user when there is information missing or a conflict within the given information in order to resolve these issues in a collaborative way. Our main requirements for this logic are:

- Expressivity:
  Whether the logic has the concepts required to express the user needs
- Formal properties:
  Whether the logic satisfies certain formal properties about the (relations between) concepts
- Understandability:
  Whether concepts in the logic are understandable and aligned with people's mental model of how they express their needs
- Adaptability:
  Whether the logic is able to handle new input and changes to existing information
- Explainability:
  Whether conclusions or conflicts of the logic can be traced to the information they originate from

The first three requirements concern the structure of the knowledge base and the language of the logic. The logic must be able to express all the concepts that are relevant for the personal agent and the needs of the user. This may for example include goals, preferences and available actions for the user and the agent. Additionally, the formal properties of the concepts and the relations between them need to be satisfied by the logic. This is of course dependent on the concepts that are included in the logic but could concern things such as making sure a preference relation has the suitable properties or that the set of all goals can be partially inconsistent. Lastly, we are aiming to make the structure of our logic understandable to the user by requiring it to align with the user's mental model of how they express their needs. If the way the logic is built and the way it reasons is closer to the way the user themselves make decisions, it may be easier for them to comprehend the conclusions of the agent.

The fourth requirement focuses on making sure that the agent can be adapted through interaction between the agent and the user. This means that when the user makes an additional input or a change to existing information, the logic is able to incorporate this, while also checking whether the update causes a conflict with the existing information and which effect it has on the conclusions of the reasoning process. We may also want to make sure that there is certain information which

cannot be changed by the user, for example to ensure functionality or safety of the agent.

Our fifth requirement is needed in order for the agent to also communicate with the user. We have mentioned that we want the agent to be able to explain how a certain conclusion was reached or why a certain update causes a conflict. For our logic that means we need to be able to trace which information has been used during the reasoning process and extract this information when necessary. We note that this is not the same thing as creating the concrete explanation that is presented to the user but rather a first selection of which information might be relevant for this explanation.

3. Related Work

Our work relates to several different areas of research, including both theoretical foundations and more practical applications.

Much of the research done on non-monotonic logic has been conducted with the intention of replicating human-reasoning [9]. In particular this refers to defeasible reasoning, which is necessary when a conclusion may need to be withdrawn if further information becomes known. This may occur if a statement is derived from incomplete information and represents the most plausible or common rather than the only possible conclusion. For example, Default Logic [10] is designed to capture the notion that a conclusion is normally true, allowing us to use it in our reasoning process if there is no argument which explicitly contradicts it. An example of this could be that if a user has an established routine, the statement "It is Wednesday" is enough to reasonably conclude that "Today the user will go swimming" is also true. However, when provided with additional information such as "The swimming pool is closed" or "The user goes for a run" which entails that assuming "Today the user will go swimming" is no longer consistent with the existing information, the conclusion is rejected. Other examples of non-monotonic reasoning techniques include the closed world assumption [11], [12], meaning that we assume the information we have about the world is complete, auto-epistemic reasoning [13], which allows an agent to reason about their own knowledge and beliefs, and logics of belief revision [14], [15] that include operations that introduce or remove belief sentences.

Within the context of artificial intelligence research these reasoning techniques were often developed to replace human reasoning and create an autonomous agent which is capable of making decisions by itself [9]. We are instead aiming to use non-monotonic logics to facilitate the process of adjusting conclusions based on new information. Additionally, there are several methods which allow us to explicitly reason with conflicts that may occur, either between multiple acceptable conclusions or between derived conclusions and certain known or desirable statements. Such techniques include adding a specificity [16] or preference relation to the logic [17] or differentiating between skeptical and credulous reasoning [18]. By being able to represent these conflicts explicitly, we aim to make it easier for the user to contribute to their resolution.

Another method which relates to our work is the use of argumentation formalisms [19] [20]. Some of these formalisms also model non-monotonic inferences,

but there there are also some which characterize argumentation as a form of dialogue. In particular, these systems model an interaction which is designed to resolve a conflict of beliefs. Examples of these types of argumentation formalisms are automatic persuasion systems [21] and dialectical argumentation [22], both of which have been studied in the context of behavior change assistants.

The importance of shared mental models in Human-AI interaction has been covered extensively in the literature, a review can be found in [23], while a conceptual analysis can be found in [24]. In order to elicit the user's mental model a variety of techniques have been studied, most commonly used are interviews with the user [25], [26]. Mental models have also been studied in psychology, which may also provide a theoretical framework to base our research on [27], [28].

The need for an explainable reasoning process has long been recognized, with [29] stating that effective explanations must be presented in form of a dialogue which allows for partial explanations and follow-up questions. In order to ensure that an agent is understandable and can clarify its output, the designers of the agent need to provide explanations in a suitable way. This combines research from psychology, education and interaction design but most importantly requires the justifications of the agent to be available. There are multiple approaches to making this information explicit in the logic. One option is to include some form of inference tracking within the language of the logic, which makes it possible to retrace the proof that the agent has created during its reasoning process. Another option is to use methods from justification logic [30] or argumentation logic [31],[32], which include the possibility to reason explicitly about arguments that support or oppose a statement. These justifications are often argued to be more intuitive than proofs as their structure is easier and they do not require knowledge of a proof system in order to be understood. An alternative approach, especially used for determining the cause of a conflict, is axiom pinpointing or finding a minimal theory. This consists of an iterative method which is used to find a minimal set of axioms or statements that entails the undesired consequence, which can provide a helpful starting point to resolving the conflict [33], [34], [35].

4. Research Questions

The overarching question that our research hopes to answer is: How can we combine techniques from non-monotonic logic and knowledge representation and reasoning to create a logical framework for a behavior support agent that enables direct interaction with the user? We can break this down into the following research questions which are based on the requirements we have mentioned in Section 2.

RQ 1 **Which structure should the logic have in order to capture relevant concepts but also align with the user's mental model?**
Answering this question will likely require some compromises to be made, especially considering that the exact mental model is dependent on the individual user. Ideally we will be able to make out a minimal structure which contains all the formal requirements for a functioning support agent which can then be personalized by the user using the methods from RQ2. This also includes deciding which techniques from non-monotonic reason-

ing and non-monotonic logic are best suited for each aspect that we want our logic to contain.

RQ 2 How can we enable the user to directly input and change information in order to adapt the agent to their needs?
A part of this question is also determining which types of updates are necessary, which will also depend on the structure we obtain from RQ1. For each different type of update we then need to assess the formal properties and which method we can use to implement them in our logic. In order to ensure that the functionality and safety of the support agent can be maintained, we also need to consider how we can test whether an update is valid and which conflicts may arise. The question of how to resolve these potential conflicts is also an element of RQ3.

RQ 3 How can we determine where in the logic conclusions and conflicts originated?
As we have seen in Section 3, there are multiple methods that have been studied in this context. Finding the solution which works best in the context of our work will depend on the results of RQ1 and RQ2, meaning the structure of our logic and the type of conclusions and conflicts we want to consider. While we will not be generating explanations in natural language, we will consider which information is relevant in each situation. For example, in some situations it may be helpful to determine possible ways to solve a conflict which can be presented as a suggestion to the user or to find the differences between two concrete situations to show the user the impact of their choices.

5. Approach

The three main research questions we have given in the previous section should not necessarily be treated as independent steps to achieve our main goal but rather as interconnected approaches which are part of an iterative process in order to create a logic which answers all three questions at once. We plan to continuously update our logic with the results obtained from our work on each of these research questions and then reevaluate whether it still aligns with the requirements we have specified earlier.

As we are aiming to creating an abstract logical framework which can be implemented in a behavior support agent, we are basing our approach on the bottom-up procedures for developing grounded theory [36]. Specifically, we will begin by following a case study approach [37] and analyzing a few selected behavior support agents for their requirements. We are interested in the concepts that are needed and the relationships between them, the mental model of the user and possible updates and conflicts that may occur while the agent is in use. Besides the formal requirements of the update itself we also need to pay attention to the requirements we want for the validity of an update. This may include specifying certain information which cannot be changed or contradicted or testing for certain conclusions before fully incorporating the update. We also analyze our use-cases to determine situations in which the user may want an additional explanation.

We will then generalize these requirements to more abstract concepts and formal properties that these should satisfy by comparing the results of each use-case analysis.

One of the use-cases we will be analyzing is the diabetes support agent that we have already mentioned. During our analysis of this agent we will first identify the relevant concepts that the agent must include such as goals, actions, preferences, medical data and critical safety values, for example. We will also take note of the relations between these concepts, like what concepts the user can express preferences between, what kinds of requirements and effects actions might have and what the critical safety values need to be compared to. Besides exploring these formal requirements for the agent, having a concrete use-case also allows us to study the user's mental model in order to understand how they structure the given situations themselves. Once we have an overview of the concepts that are relevant for the agent, we can also analyze the potential dynamics that should be incorporated. In the example of the diabetes support agent, the user may want to change their preferences regarding the types of exercise or the possible schedule, but they should not be able to delete the goal of being active completely. Adding an additional goal which partially contradicts an existing goal may be acceptable in some cases, like wanting both comfort and fitness, which cannot be fully achieved at the same time, however in other cases these conflicts should be resolved, for example if the user states they want to follow a diet which would be unhealthy for them considering their current exercise plan. Observing these types of dynamics and potential conflicts is crucial to ensuring that our logical framework will fulfill the necessary requirements in order to be effective and safe when implemented into an agent. We have chosen the diabetes use-case in particular because this application is already being researched [38], allowing us to include interviews with domain experts and existing data in our analysis.

As well as studying the user's mental models of our example applications, we will also be conducting a literature review to survey the existing theoretical work on this topic. While each individual user's mental model may vary, using existing frameworks can ensure that our work is also based on established results.

In parallel to this, we intend to survey the literature on techniques from knowledge representation and non-monotonic reasoning. The purpose of this is to gain an overview of the formal properties of these methods as well as their potential strengths and weaknesses for possible applications. Additionally, we need to study what effect each update has on the knowledge base as a whole and where conflicts may arise. For example, default logics as we have described in Section 3 are useful for describing how things normally happen, which might make them suitable for capturing a user's routine. Additionally, by adding specificity relations to these logics we can easily include exceptions to these routines. However, by allowing the user to add or remove default rules as they wish, it may be possible that the agent cannot deduce any possible suggestions anymore, which would impact its effectivity.

We will also conduct a literature survey on existing methods which offer insight into where conclusions originate such as the ones described in Section 3. We can then compare the requirements we have compiled with the capabilities of the logics we have studied in order to determine a suitable base logic as our

starting point. This can then be expanded upon by using other techniques in order to fulfill more of our requirements. This procedure is similar for each of our research questions and we believe that a flexible, possibly iterative approach is beneficial for our research. While we may find that certain techniques are best suited to model the general structure of our logic it is possible that these do not allow the adaptability or explainability we require. This will require us to find a compromise which works best in the context of our use-case.

The use-cases we analyze at the beginning also give us an opportunity to evaluate our work throughout the course of our research, by comparing the capabilities of our logic to the requirements of our examples.

6. Evaluation

Throughout our research we will be evaluating our logic by applying it to the selected use-cases that we will choose at the beginning of our research. This is intended to keep our work grounded in realistic applications but it is also helpful as a frame of reference to compare the capabilities of our logic to. Since we plan to choose use-cases which are already being studied independently of our work, this also gives us the opportunity to ask experts to validate certain aspects of our logic, such as whether it aligns with their user models or whether it is adequately expressive.

In order to evaluate whether our results can also be generalized to other applications, we will also use our logic to model a new use-case which has not been considered during our work. This may also give us the opportunity to conduct a user study in order to test how well the requirements we have given in Section 2 are satisfied when the agent is in use.

Besides the application, we will also evaluate our work theoretically. The logic itself should be proven to satisfy all formal requirements such as the reasoning process being sound and the concepts satisfying the necessary properties. The structure of the logic should align with results from user modeling and psychology.

7. Conclusion

With this research we intend to provide a logical framework which enables direct interaction between a personal support agent and the user. We aim for our work to result in general methods which can be used to make an agent adaptable to the user's flexible needs by allowing the user to directly influence the knowledge base of the agent. Additionally, the logic will be designed to facilitate creating an agent which is understandable to the user, both by having the structure align with the user's model and by providing the information needed to explain the agent's reasoning process.

We expect the general methods we obtain to be sufficiently flexible for them to be adapted depending on the specific intention and context of an application. This also means that our research can be extended in multiple directions in order to study how the tools that our framework offers can be used optimally. For

example, in our work we assume that we receive clear information from the user in order to update the knowledge base. However, the question of how the user interface of the agent can enable this input or how the relevant information can be extracted from these inputs depends on the setting that the agent operates in. Similarly, we only aim to extract information from the logic which can be used to explain the reasoning process. In order to ensure that the user is able to understand how the agent has reached a certain conclusion, this information will likely have to be filtered further and presented in an appropriate way.

Funding

This research was funded by the Hybrid Intelligence Center, a 10-year programme funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research, https://hybrid-intelligence-centre.nl, grant number 024.004.022.

Acknowledgments

I would like to thank my supervisors Prof. Dr. Dirk Heylen, Dr. Birna van Riemsdijk and Dr. Victor de Boer for their support and advice.

References

[1] Oinas-Kukkonen H. Behavior Change Support Systems: A Research Model and Agenda. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Persuasive Technology. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 4-14. doi:10.1007/978-3-642-13226-1_3.

[2] van Riemsdijk MB, Jonker CM, Lesser V. Creating Socially Adaptive Electronic Partners: Interaction, Reasoning and Ethical Challenges. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems. AAMAS '15. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems; 2015. p. 1201–1206.

[3] Sundar SS. Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). Journal of Computer-Mediated Communication. 2020 01;25(1):74-88. doi:10.1093/jcmc/zmz026.

[4] Akata Z, Balliet D, de Rijke M, Dignum F, Dignum V, Eiben G, et al. A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. Computer. 2020;53(8):18-28. doi:10.1109/MC.2020.2996587.

[5] Amershi S, Inkpen K, Teevan J, Kikin-Gil R, Horvitz E, Weld D, et al. Guidelines for Human-AI Interaction. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; 2019. p. 1-13. doi:10.1145/3290605.3300233.

[6] Anjomshoae S, Najjar A, Calvaresi D, Främling K. Explainable Agents and Robots: Results from a Systematic Literature Review. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems. AAMAS '19. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems; 2019. p. 1078–1088. doi:10.5555/3306127.3331806.

[7] Haque AB, Islam AKMN, Mikalef P. Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. Technological Forecasting and Social Change. 2023;186. doi:10.1016/j.techfore.2022.122120.

[8] Lockey S, Gillespie N, Holm D, Asadi Someh I. A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions. In: Proceedings of the 54th Hawaii International Conference on System Sciences; 2021. p. 5463 -5472. doi:10.24251/HICSS.2021.664.

[9] McCarthy J, Hayes PJ. Some Philosophical Problems from the Standpoint of Artificial Intelligence. In: Meltzer B, Michie D, editors. Machine Intelligence 4. Edinburgh University Press; 1969. p. 463-502.

[10] Reiter R. A logic for default reasoning. Artificial Intelligence. 1980;13(1):81-132. Special Issue on Non-Monotonic Logic. doi:10.1016/0004-3702(80)90014-4.

[11] McCarthy J. Circumscription - A Form of Non-Monotonic Reasoning. Artificial Intelligence. 1980;13:27-39. doi:10.1016/0004-3702(80)90011-9.

[12] Pratt I. Closed World Assumptions. In: Artificial Intelligence. Red Globe Press London; 1994. p. 65-84. doi:10.1007/978-1-349-13277-5_4.

[13] Moore RC. Semantical considerations on nonmonotonic logic. Artificial Intelligence. 1985;25(1):75-94. doi:10.1016/0004-3702(85)90042-6.

[14] Hansson SO. A Textbook of Belief Dynamics: Theory Change and Database Updating. Dordrecht and Boston: Kluwer Academic Publishers; 1999.

[15] HANSSON S. New operators for theory change. Theoria. 1989;55(2):114-32. doi:https://doi.org/10.1111/j.1755-2567.1989.tb00725.x.

[16] Rintanen J. On Specificity in Default Logic. In: IJCAI'95: Proceedings of the 14th international joint conference on Artificial intelligence; 1995. p. 1474–1479. doi:10.5555/1643031.1643090.

[17] Langholm T. Default Logics with Preference Order: Principles and Characterisations. In: Cervesato I, Veith H, Voronkov A, editors. Logic for Programming, Artificial Intelligence, and Reasoning. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 406 -420. doi:10.1007/978-3-540-89439-1_29.

[18] Horty J. Skepticism and Floating Conclusions. Artificial Intelligence. 2002 02;135:55-72. doi:10.1016/S0004-3702(01)00160-6.

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

[20] Prakken H. In: Baroni P, Gabbay D, Giacomin M, van der Torre L, editors. Historical overview of formal argumentation. vol. 1. College Publications; 2018. p. 73-141.

[21] Hunter A. Invited Talk: Computational Persuasion with Applications in Behaviour Change. In: Arai S, Kojima K, Mineshima K, Bekki D, Satoh K, Ohta Y, editors. New Frontiers in Artificial Intelligence. Cham: Springer International Publishing; 2018. p. 336-6.

[22] GRASSO F, CAWSEY A, JONES R. Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. International Journal of Human-Computer Studies. 2000;53(6):1077-115. doi:https://doi.org/10.1006/ijhc.2000.0429.

[23] Andrews R, Lilly J, Srivastava D, Feigh K. The role of shared mental models in human-AI teams: a theoretical review. Theoretical Issues in Ergonomics Science. 2022 04;24:1-47. doi:10.1080/1463922X.2022.2061080.

[24] Jonker C, Riemsdijk M, Vermeulen B. Shared Mental Models - A Conceptual Analysis. In: Coordination, Organizations, Institutions, and Norms in Agent Systems VI; 2010. p. 132 -151. doi:10.1007/978-3-642-21268-0_8.

[25] Jones N, Ross H, Lynam T, Perez P. Eliciting Mental Models: a Comparison of Interview Procedures in the Context of Natural Resource Management. Ecology and Society. 2014 02;19:13. doi:10.5751/ES-06248-190113.

[26] Memon T, Lu J, Hussain FK. An Enhanced Mental Model Elicitation Technique to Improve Mental Model Accuracy. In: Lee M, Hirose A, Hou ZG, Kil RM, editors. Neural Information Processing. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 82 -89. doi:10.1007/978-3-642-42054-2_11.

[27] Rouse W, Morris N. On Looking Into the Black Box. Prospects and Limits in the Search for Mental Models. Psychological Bulletin. 1984 10;100. doi:10.1037/0033-2909.100.3.349.

[28] Johnson-Laird PN. Mental models in cognitive science. Cognitive Science. 1980;4(1):71-

115. doi:10.1016/S0364-0213(81)80005-5.
[29] Moore J, Swartout W. A Reactive Approach to Explanation: Taking the User's Feedback into Account. In: Natural Language Generation in Artificial Intelligence and Computational Linguistics; 1989. p. 1504 -1510. doi:10.1007/978-1-4757-5945-7_1.
[30] Artemov S. The logic of justification. The Review of Symbolic Logic. 2008 12;1:477 513. doi:10.1017/S1755020308090060.
[31] Liao B, Anderson M, Anderson S. Representation, justification, and explanation in a value-driven agent: an argumentation-based approach. AI and Ethics. 2020 09;1. doi:10.1007/s43681-020-00001-8.
[32] Cyras K, Fan X, Schulz C, Toni F. Assumption-Based Argumentation: Disputes, Explanations, Preferences. IFCoLog Journal of Logics and Their Applications. 2018 02;4:2407 -2456.
[33] Peñaloza R. Explaining Axiom Pinpointing. In: Lutz C, Sattler U, Tinelli C, Turhan AY, Wolter F, editors. Description Logic, Theory Combination, and All That: Essays Dedicated to Franz Baader on the Occasion of His 60th Birthday. Cham: Springer International Publishing; 2019. p. 475 -496. doi:10.1007/978-3-030-22102-7_22.
[34] Dev Gupta S, Genc B, O'Sullivan B. Explanation in Constraint Satisfaction: A Survey. In: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence; 2021. p. 4400 -4407. doi:10.24963/ijcai.2021/601.
[35] Bogaerts B, Gamba E, Guns T. A framework for step-wise explaining how to solve constraint satisfaction problems. Artificial Intelligence. 2021;300:103550. doi:10.1016/j.artint.2021.103550.
[36] Corbin J, Strauss A. Grounded Theory Research: Procedures, Canons and Evaluative Criteria. ZfS - Zeitschrift für Soziologie; ZfS, Jg 19, Heft 6 (1990); 418-427. 1990 03;13. doi:10.1007/BF00988593.
[37] Slight S, Cresswell K, Robertson A, Huby G, Avery T, Sheikh SA. The Case Study Approach. BMC medical research methodology. 2011 06;11:100. doi:10.1186/1471-2288-11-100.
[38] Harakeh Z, de Hoogh IM, van Keulen H, Kalkman G, van Someren E, van Empelen P, et al. 360° Diagnostic Tool to Personalize Lifestyle Advice in Primary Care for People With Type 2 Diabetes: Development and Usability Study. JMIR Form Res. 2023 Mar;7. doi:10.2196/37305.