Help me make a dinner! Challenges when assisting humans in action planning

Gregor Behnke[1], Benedikt Leichtmann[3], Pascal Bercher[1], Daniel Höller[1], Verena Nitsch[3], Martin Baumann[2], and Susanne Biundo[1]

[1] *Institute of Artificial Intelligence, Ulm University, Ulm, Germany*

[2] *Institute of Psychology and Education, Department Human Factors, Ulm University, Ulm, Germany*

[3] *Human Factors Institute, Universität der Bundeswehr München, Neubiberg, Germany*

***Abstract*—A promising field of application for cognitive technical systems is *individualised user assistance* for complex tasks. Here, a companion system usually uses an AI planner to solve the underlying combinatorial problem. Often, the use of a bare black-box planning system is not sufficient to provide *individualised* assistance, but instead the user has to be able to control the process that generates the presented advice. Such an integration guarantees that the user will be satisfied with the assistance s/he is given, trust the advice more, and is thus more likely to follow it. In this paper, we provide a general theoretical view on this process, called mixed-initiative planning, and derive several research challenges from it.**

1. Introduction

The ability to provide correct and individualised advice to a human user is one of the core capabilities a *companion system* should have [1], [2]. To provide this assistance – especially in complex situations, a companion system can use an AI planning system. It is able to provide a course of action that leads to the goal the user desires. That way, it provides the companion system with the capabilities needed to help the user. One example of such a system is an assistant which supports users in setting up a home theatre system [3].

However, in many situations, simple advice generated by an AI planner is not sufficient. E.g. when the user has individual preferences unknown to the planner or when s/he herself/himself wants to be "in control" of the steps s/he has to execute. In the extreme, an AI planner should not make live-or-death decisions, e.g., when planning systems are applied in a military context [4]. However, standard planning systems do not allow for a user to take control of the planning process. To solve this problem, the idea of mixed-initiative planning was developed. Here planning is not seen as a mere optimisation process, but as a dialogue between the user and the planner. Such planners have been developed and deployed in the past in several application scenarios [5], [6], [7], [8], [9]. However, all these systems were developed for a single application and none of them provides a framework for mixed-initiative planning in a general, domain-independent setting. So far, there has not been any coherent research on the theory and application of mixed-initiative planning. In this paper, we discuss what the aims and objectives of a mixed-initiative planning system are and which research challenges result from them. We first discuss relevant psychological concepts and then introduce the challenges we think should be addressed in future research on mixed-initiative planning.

2. Relevant Principles of Cognition

Human planning behavior is dynamic. Although past studies of human planning behavior indicate that humans take a hierarchical, top-down, total order planning approach (see [10], [11]), the problems that were used in these studies have been rather abstract and well-defined puzzles and not situated in real-world settings [12]. When focusing on real-world scenarios with ill-defined problems, human planning behavior is rather opportunistic, non-hierarchical, bottom-up and in partial order [12], [13]. Humans do not have a full representation of the problem space, all possible actions, initial state or final state, so they use several heuristics in order to arrive at a solution. At the beginning, a prototype plan based on goals and on identified key steps as a result of environmental monitoring is generated. This prototype plan is changed, replaced by alternative back-up plans and procedurally tailored to the situation in phases of plan projection and revision. During implementation, the plan is periodically reevaluated, adjusted and transformed to cope with environmental changes identified by marker events [14]. Planning is thus highly dynamic and influenced by several interpersonal and situational factors.

Humans have limited resources. One aspect why planning behavior appears to be non-hierarchical, bottom-up, and in partial order is because human beings are limited in their resources. Classical studies mention seven plus / minus two chunks [15] or even just four chunks [16] that people can use to store and manipulate information in their working memory. There is no doubt that planning - especially for ill-defined problems in real world scenarios - is a demanding, resource-intensive activity [14], so people have to overcome their limits in cognitive resources by using heuristics and other strategies to lower working memory load. For example, humans use the heuristic of analogy [17], [18], [19] or a heuristic similar to hill-climbing, whereby problems are decomposed into sub-problems with sub-goals, and satisfying moves that promise progress in improving the current state against the hypothesized goal properties are selected [20]. In real-world diagnostic reasoning, for example when a physician tries to find an explanation of the patient's symptoms, the complexity of the problem is greatly reduced by the use of context information. One major source of context information is the reasoner's current explanation [21], [22]. That is, from all possible explanations for a new observation very often only those are considered that are compatible with the current explanation. In routine problems, such as the diagnosis of well-known diseases, the pre-selection of possible candidates for the problem solution and the interpretation of new observations is based on automatic memory processes [23]. Of special importance here is the usage of long-term memory and more elaborated knowledge structures by experts in such routine problems. The theory about situation comprehension by Durso, Rawson & Girrotto (2007) can explain the processes that result in the better performance of experts [24]. By these implicit processes that interpretation of a new observation gets activated most and consequently selected that is most compatible with the current explanation of the previous observations. Whereas such processes can significantly limit the number of possible explanations to consider it can also blind the reasoner for possible alternative explanations and therefore can lead to suboptimal performance. Hence, if someone wants to understand human planning and problem-solving behavior, one has to inevitably understand how various aspects of human cognition, such as automatic involuntary memory processes interact with deliberate reasoning processes [25].

Cognition is situated. Besides individual differences in cognitive resources, differences of situational factors also influence human planning behavior. From situated cognition perspective [26] problem solving and planning is not seen as an abstraction with formal structure that is the same across different settings. According to Kirsh (2009), problems are tied to a specific situation and planning takes place by a situation-specific reasoning processes deeply embedded in the context. This perspective makes it hard to generate hypotheses about planning which are true for several situations, or to extract common aspects of planning except concepts of general psychology such as memory. Evidence for this view are studies showing that solving the same mathematical problems differs depending on the setting, for example when the problem is framed differently (e.g. [27]). Another study by Ormerod & Ball (1993) shows that experts with a similar level of expertise (and thus domain-specific knowledge) generate slightly qualitatively different solutions depending on their background [28]. The process of planning will thus differ radically between individuals [29], as the employed planning strategy can vary depending on individual experience and situational factors.

Aspects of human-machine interaction. As planning is a resource-intensive process, intelligent support systems, which are less limited in their resources and can choose amongst solutions in a more systematic way, can provide considerable help for humans in tasks that require planning. Such systems have to meet certain criteria of human-machine interaction design to provide optimal support and to ensure that users can benefit from the full potential such systems can offer. According to Christoffersen & Woods [30], automated systems in general have to fulfill two criteria to act as a real companion for the user: They have to be observable and directable. Observability opens up the black box of the automated system enabling the user to observe and understand the system's environment perception, its state, actions, and plans. Observability is the prerequisite for the construction of a shared situation representation between the human operator and the automated system, which in turn is vital for a system to be predictable. Directability means that the operator is able to easily change the system's priorities and plans in the face of changed conditions. To sum up, for an intelligent support system to be most effective it has to be designed as a team player for the human partner that is observable and directable.

3. Challenges

Given these principles of human cognition relevant for planning behaviour in this section we discuss four of the – from our point of view – most relevant questions and challenges that arise when humans and AI planning systems are working together. To develop these challenges, we first have to define what the objective of a joint human/computer planning process is. In the context of companion systems [1], [2], such a cooperation aims at selecting a course of action that both leads to the goal the user desires and satisfies all other constraints and preferences s/he might have – in other words: it takes the user's cognitive and personality characteristics into account. To show the contrast between these two objectives (and as a running example throughout this section), we will assume the following example. The user's current objective is to cook a complex, three-course menu in a given time frame and s/he wants to plan the steps necessary for preparing the meal aided by an AI planner. To satisfy the goal, the planner could produce any sequence of actions that results in the desired menu, i.e., the planner has to determine which preparatory steps are necessary and in which order they have to be executed to finish the meal in time. However, there can be courses of action that may cause inconvenience to the user, e.g., s/he would have to perform steps s/he is unaccustomed to.

If we assume that the planner's domain model is a correct abstraction of the world for the purposes of the planning task, the system is left to determine the goal of the user and to determine her/his preferences. If it knew both, it could determine the optimal (in terms of user preferences) plan with high accuracy and present it to the user. Since in general we cannot assume to know the preferences of a single, specific user *a priori*, the system has to determine the preferences of the single user, a process called preference elicitation [31]. Classical methods for preference elicitation are based on an expensive and lengthy questioning of the user on potential preferences until a single unique preference function has been determined. Such an interview-style interaction process is not appropriate when the user is expecting to be assisted in solving a complex task by an AI planner, e.g., because the elicitation process would not show any progress in planning. Since *companion systems* should be immediately responsive, a lengthy testing phase prior to actually using the system might cause users to abort the planning process altogether, which should be averted at any cost. Also the interpretation of preferences itself can become problematic, if one wanted to use preference elicitation directly. For example, the user herself/himself may not have a well-defined mental model with respect to which s/he can answer questions, e.g., her/his preferences may contradict themselves. Furthermore human planning processes are highly dynamic, so preferences might change because of the opportunistic fashion of the planning process and some preferences might come up right during the planning process itself [12], [13].

To determine these preferences nonetheless, the user can be directly integrated into the planning process. The aim of this integration is to let the user actively participate in the planning process, while simultaneously being able to observe how s/he makes planning decisions and which instructions s/he gives to the planner. From this interaction, the planner can infer the user's preference model and – if confident enough in its determination – perform tasks during the planning process itself. Ideally, this planning process ends with the solution that the planner would choose based on the determined preferences, ensuring that the user perceives that s/he has reached the plan together with the planner and not that it was dictated to her/him. In this case the system meets the criterion of directability [30]: the user is able to adapt the planner's activities to her/his needs and based on changing priorities. Additionally, the system's processes should be observable for the user. Consider that a system might, e.g., know a certain user prefers to use public transportation, the system's plan includes using a car because public transportation is not available. This difference between system's and user's plans has to be made observable and the system has to provide explanations for this difference to achieve the user's acceptance.

Clearly, this abstract view on integrating the user is not sufficient to construct a mixed-initiative planning system, but based on this general view on the process, we can identify four challenges that have to be addressed.

3.1. Challenge 1: What is the topic of discourse?

The most fundamental question in designing a mixed-initiative planning system is to choose the topic of the interaction with the user. In most existing mixed-initiative planning systems, the interaction focusses on plans. Here, the user is repeatedly shown a plan (which could be both a solution or an yet unfinished plan), for which s/he is either presented options to alter it or can freely suggest modifications. The planner then takes these changes under consideration and shows the user an accordingly adapted plan, starting the process again. In this paper, we assume that the mixed-initiative planning system will use this scheme.

However, other means to interact with the user can be imagined, which also let the user think that s/he is taking part in the planning process and is making progress, while simultaneously eliciting her/his preferences. For example, the system could ask direct questions about preferences (against which we have argued before): "When going to the supermarket, do you prefer riding a bicycle or driving?" Similarly, one may also think of a more reactive interaction, in which the system asks the planner what to do next (i.e. eliciting opportunistic planning on the user's part) and the planner only observes the user in creating a plan by herself/himself. However, this scheme of interaction is not well suited to assist the user, as the planner does not provide substantial support in any meaningful way.

Even when we restrict the interaction with the user to questioning her/him about a current plan, there are still design-choices remaining. Most prominently, one has to divide what kinds of plans should be presented to the user. Current mixed-initiative planning systems present only solutions to the user, i.e., plans which already fulfil the goal and the preferences the user has stated. This can pose severe problems when interacting with the user, e.g., because these plans can be extremely complex and might overwhelm users because of the human's limited cognitive resources (see Section 3.2. of this paper). In a general mixed-initiative framework, the planner should be able to interact with the user also based on partial plans, which are not yet solutions. This also takes into account that human users do not have a full representation of the problem and tend to decompose problems into sub-problems in order to overcome their limits [20]. But even if we decide to present partial plans, there is still the question of which degree of detail is needed. The appropriate degree of detail might vary depending on interindividual differences for example in expertise/background knowledge or in cognitive abilities.

Also, by presenting partial plans, the user can control the process of refining it into a solution, and exert the maximum amount of control over the planning process which again fulfils the criterion of directability of the system in an cooperative human-machine system. This general process seems to be well suited for the sought-for integration of users into the planning process. Especially, it also aligns well with the way humans tend to perform planning tasks.

Whenever plans are presented that are not yet solutions, another question arises: Is the planner allowed to present plans to the user, which cannot be refined into a solution any more (a so-called *dead end*)? If the whole space of partial plans was known, we could determine whether a partial plan can possibly be refined into a solution. In practice, we cannot always make this determination – in sufficient time or at all. Showing the user a *dead end* plan becomes an issue either if the user tries to refine the plan into a solution and fails, or if the flaw in the plan is apparent enough such that the user can detect the problem himself/herself. In such situations, the user's trust in the system can be severely harmed, but we cannot always guarantee to filter all dead end plans. So the planner has to be able to cope with the situation that a presented plan turns out to be a dead end.

3.2. Challenge 2: How to react to the user?

The next challenge is in some sense the counterpart to the previous one: How should the system react to the user's utterances, i.e., which plan should be shown next?

If we take an abstract view on the planning process, it is the process of jointly navigating through the space of all possible plans. One might think that the user's choices and instructions determine a unique path through that space, which is unfortunately often not the case. The plan shown after a user utterance should be a result of the instructions s/he has just given and not just any arbitrary plan, showing which might be helpful for eliciting the user's preferences. In many circumstances, there is not one single plan that results from such instructions, but rather a set of possible alterations, hence the question arises which plan to choose. Take for example the above-mentioned example to cook a three-course meal and imagine that the planner proposes to cook a cream-based sauce. Here the user might not want to make such a sauce, leaving the planner with a variety of options. For example, we could make a béchamel- or a fond-based sauce instead, where the latter would also entail to interrupt the cooking process in order to buy fond in the supermarket. Here the planner has to select one of the plans to present it to the user. It can either choose the one more similar to the original plan (i.e. making béchamel sauce), or the "easier solution" – the plan with less or less complex actions (i.e. just buy a fond). If we take the second option, the resulting plan is not even a direct neighbour of the original plan in the space of all plans, as there is no single refinement or alteration that transforms them into each other. There is no a priori correct answer, as we should keep consecutively shown plans as similar as possible to not overexert the user's working memory, but also have to show a plan and alteration that the user can easily understand and will be content with. A mixed-initiative planning system has to provide an answer to this question, which should weight these desiderata. Therefore, an understanding of human cognition is needed. Not only the limited human resources, e.g., in terms of working memory capacity, need to be considered, but also the current situational representation of the problem. These structural differences between representations in form of mental models depending for example on expertise level or the individual's background cause differences in our usage of problem solving strategy or domain-specific reasoning processes [28], [29]. The impact of those interindividual differences in expertise and cognitive abilities or personality factors on joint human-machine problem solving and planning is still not clear.

One also has to be aware of the problem that a preference once uttered by the user might become irrelevant during the planning process. The user might state that s/he prefers to use béchamel sauce instead of fond, but in the final plan, the course containing the sauce has been replaced with a soup, making the initial question concerning the sauce superfluous. This situation can be very irritating to the user, as it would have been possible to find a plan without ever asking. A solution has to be found to either avert this situation altogether, minimise its impact, or to be able to explain the user this situation appropriately to support the creation of a shared situation representation.

This consideration is also relevant when considering how to change a plan if the user requests it. In the sauce example, the user might have instructed to simply replace the cream sauce with making a fond sauce, which however would have made it impossible to finish the menu in time (making fresh fond takes time). If the planner would blindly follow the instruction, it would not provide assistance to the user as it would propose a wrong course of action. A successful mixed-initiative planner therefore should also perform the necessary changes in this situation that ensure that the resulting plan can still be refined into a solution (i.e. buying the fond instead of making it). The challenge thus is how to determine these necessary changes, and to select those that are appropriate.

3.3. Challenge 3: Inferring the model

In a mixed-initiative planning system, we have to answer the question of how to use the information we got from the user to find a plan best suited for her/him. Since we have to determine the user's characteristics in terms of personality and cognitive-psychological aspects to do this, the challenge here is to extract these aspects from the user's utterances.

One of the most apparent problems is that the user's utterances typically do not have a single valid interpretation. For example, if s/he instructs the planner to find a plan "without the cream sauce", does this mean that s/he *never* wants to use cream sauces or that s/he does not want to use it in the *the current situation*. One of the main questions in this case is to determine whether the preferences are the result of situational factors or the result of psychological traits and general constructs. The central technical question is how the planner should interpret the user's preferences internally. In the past, there have been many approaches to incorporate some kinds of preferences and constraints into the planning process. These include action costs, action preferences, soft-goals, or incorporating LTL constraints. Here the designer of the system has to make the decision which formalisms to support. This choice also influences the resolution of ambiguous user inputs, as the system will typically interpret them in a way it can handle (e.g. if the system can only handle LTL, it interprets all constraints as hard ones).

We might also incur the problem of oversubscription – which is that there is no plan that satisfies all posed constraints. This will happen especially if the plan is changed based on the request of the user multiple times while the user is changing her/his opinion. For example, s/he first states that s/he wants to make a fond-based sauce, then sees the consequences of this choice (driving to the supermarket), and subsequently decides to direct the planner to use the béchamel sauce. In this situation – and maybe in others and more complex ones – it is not possible to satisfy all of the user's preferences simultaneously, but the planner has to proceed nevertheless. As such, s/he has to either ask its user for help ("What constraint shall I drop?") or determine a constraint to relax or drop. In both cases, the planner has to first detect the oversubscription problem and subsequently to determine a correct course of action which is acceptable to the user because of observability and directability of it's processes in this situation. Both problems are open questions for future research. Smith [32] raised a similar concern in the context of mission planning for space flights.

So far, we have assumed that the planner knows the overall goal that the user wants to achieve. This might not be true, especially if the user uses the mixed-initiative planner to explore how difficult it will be to achieve certain goals compared to each other or together. Similarly, the user may also not yet know this goal in detail, but rather a more abstract description of it (e.g. "I want a tasty menu."). Thus the system has also to be able to work without a concrete goal (or even without any goal) and refine (or determine) the actual goal based on the given instructions.

Lastly, we may also pose the question of which planning formalism and language to use. There are two general types of planning formalisms: non-hierarchical and hierarchical ones. Beside a fine-grained advice on how to achieve a goal, hierarchical formalisms also provide a more abstract description on what needs to be done. That way, it naturally enables the communication on different layers of abstraction. It also enables the description of more complex behaviour [33], but to the cost of a higher complexity of finding a solution or changing a plan [34].

3.4. Challenge 4: Establishing shared representation through explanations

Unfortunately, the general strategy of presenting plans to the user and inquiring their opinion is often not fully sufficient to ensure that users can be effectively assisted by the planning system. One of the most important objectives in humans-computer interaction is to ensure that the user trusts the computer s/he is interacting with. If this trust is lost, the interaction will suffer or it will even be aborted by the user [35], [36]. In our scenario, the lack of sufficient understanding of the planner's behaviour can lead to such mistrust ([37], for a more general setting see [38]).

A lack of understanding can be effectively mitigated by explaining the system's behaviour to the user [39]. For planning-based systems in particular, explanations can improve the user's trust in the generated solution [7], [40]. A need for explanations was also highlighted by Smith [32], arguing that mixed-initiative planning systems need to be able to explain the choices they make. By the ability to explain plans, the system can support a shared understanding and is thus observable in it's plan generation.

In Chapter 3.2, we have argued that whenever a currently shown plan is to be altered based on the instructions of the user, the resulting plan should be chosen, such that the user's mental capacity will not be overexerted. Achieving this objective optimally is difficult, as the requested changes may lead to consecutive changes to the plan. As these changes can become complex, it is paramount to be able to explain to the user the current plan, its internal causality, and the reasons why the planner has produced this plan. These explanations can vary in the degree of detail and in the used domain-specific vocabulary. Thus the system needs to take into account user aspects such as domain-specific knowledge/expertise. E.g. novices have no elaborated mental model of the problem and thus no detailed representation of the current situation. For them, a change in the plan has to be explained in much more detail to ensure understanding as it would be the case with an expert user. Experts have much more elaborated knowledge structures they can use and thus a much more detailed representation of the situation. Such expert users might be annoyed if a system explains changes in a plan in detail, because experts are able to understand those changes on a much more abstract level. This can be explained by the theory of situation comprehension by Durso, Rawson & Girrotto [24].

There has been only little work on generating explanations for plans. There are two approaches that cover contrary questions on a plan. Seegebarth et al. presented a technique which allows for explaining *why an action is part of a plan* [41]. Their formalism can also answer questions regarding the order between actions and parameters chosen for actions. Gödelbecker et al. developed a scheme to explain *why a given planning task is not solvable*, thereby enabling to explain the user why a certain choice cannot be made [42]. However, both these approaches can have drawbacks when interacting with users. Seegebarth et at.'s work can only explain causality within a plan, i.e., it cannot explain why an action cannot be circumvented. Similarly, Gödelbecker et al.'s work is limited to classical planning and its explanations are limited to so-called "excuses" – alterations to the initial state that would make the problem solvable. These explanations are not able to give the user the necessary insights into the interdependency of actions.

We argue that extending the existing capabilities of plan-explanations is paramount for a mixed-initiative planning and should be the target of future research. These extensions should (at least) include the ability to explain *why an action cannot be avoided*, *why an action cannot be used*, and *why a certain goal cannot be achieved*. A mixed-initiative planning system should also be able to answer hypothetical questions, such as *what would happen if we replaced this action with that one*? Lastly, there should also be the option to explain how the current plan was created, especially after the planner has performed consecutive changes after a change request. Such an explanation would, e.g., answer the questions *why was this action inserted into the plan* or *why was this action removed from the plan* and take user characteristics, such as the individual's level of expertise, into account. This directive of research is also spearheaded by Smith [32].

4. Conclusion

In this paper we made the first steps towards systematic theoretical and practical research into the development of mixed-initiative planning systems. We gave a general interpretation of the interaction between user and planner as a process aimed at determining the preferences of the user in the planning problem at hand. From this interpretation, we derived four challenges that need to be addressed by a mixed-initiative planning system. Two of them focus on the way the interaction with the user should be shaped. The third challenge centers around the question how a model of the user's wishes can be derived from that interaction. The last challenge asks how the internal models of both user and planner during the interaction can be kept coherent using explanations. We think that these challenges provide an interesting and relevant direction for further research.

Acknowledgments

We acknowledge the support of the Transregional Collaborative Research Center SFB/TRR 62 "A Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

References

[1] S. Biundo and A. Wendemuth, "Companion-technology for cognitive technical systems," *KI*, vol. 30, no. 1, pp. 71–75, 2016.

[2] S. Biundo, D. Höller, B. Schattenberg, and P. Bercher, "Companion-technology: An overview," *KI*, vol. 30, no. 1, pp. 11–20, 2016.

[3] P. Bercher, F. Richter, T. Hörnle, T. Geier, D. Höller, G. Behnke, F. Nothdurft, F. Honold, W. Minker, M. Weber, and S. Biundo, "A planning-based assistance system for setting up a home theater," in *Proc. of AAAI*, 2015.

[4] K. Myers, W. Tyson, M. Wolverton, P. Jarvis, T. Lee, and M. desJardins, "PASSAT: A user-centric planning framework," in *In Proc. of the 3rd Int. NASA Work. on Plann. and Sched. for Space*, 2002.

[5] G. Ferguson, J. Allen, and B. Miller, "TRAINS-95: towards a mixed-initiative planning assistant," in *Proc. of AIPS*, 1996.

[6] J. Bresina, A. Jónsson, P. Morris, and K. Rajan, "Activity planning for the mars exploration rovers," in *Proc. of ICAPS*, 2005.

[7] P. Bercher and D. Höller, "Interview with David E. Smith," *Künstliche Intelligenz*, vol. 30, no. 1, pp. 101–105, 2016, Special Issue on Companion Technologies.

[8] J. Fernández-Olivares, L. Castillo, Ó. García-Pérez, and F. Palao, "Bringing users and planning technology together. Experiences in SIADEX," in *Proc. of ICAPS*, 2006.

[9] K. Myers, P. Jarvis, W. Tyson, and M. Wolverton, "A mixed-initiative framework for robust plan sketching," in *Proc. of ICAPS*, 2003.

[10] H. Simon, *Handbook of learning & cognitive processes: V. Human information*, 1978, ch. Information-processing theory of human problem solving, pp. 271–295.

[11] A. Newell and H. Simon, *Human problem solving*, 1972.

[12] S. Davies, *The cognitive psychology of planning*, 2005, ch. Planning and problem solving in well-defined domains, pp. 35–52.

[13] B. Hayes-Roth and F. Hayes-Roth, "A cognitive model of planning," *Cognitive science*, vol. 46, pp. 77–101, 1979.

[14] M. Mumford, R. Schultz, and J. V. Doorn, "Performance in planning: Processes, requirements, and errors," *Review of General Psychology*, vol. 5, no. 3, pp. 213–240, 2001.

[15] R. Atkinson and R. Shiffrin, *The psychology of learning and motivation*, 1968, ch. Human memory: A proposed system and its control processes, pp. 89–195.

[16] D. Broadbent, *Studies in long-term memory*, 1975, ch. The magic number seven after fifteen years, pp. 3–18.

[17] T. Ormerod, *The cognitive psychology of planning*, 2005, ch. Planning and ill-defined problems, pp. 53–70.

[18] M. Gick and K. Holyoak, "Analogical problem solving," *Cognitive psychology*, vol. 12, pp. 306–355, 1980.

[19] K. Kurtz and J. Loewenstein, "Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one," *Memory & Cognition*, vol. 35, no. 2, pp. 334–341, 2007.

[20] J. MacGregor, T. Ormerod, and E. Chronicle, "Information-processing and insight: A process model of performance on the nine-dot and related problems," *Journal of experimental psychology: Learning, memory and cognition*, vol. 27, pp. 176–201, 2001.

[21] A. Elstein, L. Shulman, and S. Sprafka, *Medical problem solving: An analysis of clinical reasoning*, 1978.

[22] T. Johnson and J. Krems, "Use of current explanations in multicausal abductive reasoning," *Cognitive science*, vol. 25, no. 6, pp. 903–939, 2001.

[23] M. Baumann, K. Mehlhorn, and F. Bocklisch, "The activation of hypotheses during abductive reasoning." in *Proc. of the 29th Ann. Conf. of the Cognitive Science Soc.*, 2007.

[24] S. G. F. Durso, K. Rawson, *Handbook of Applied Cognition*, 2007, ch. Comprehension and Situation Awareness, pp. 163–194.

[25] G. Ward and R. Morris, *The cognitive psychology of planning*, 2005, ch. Introduction to the psychology of planning, pp. 1–34.

[26] D. Kirsh, *The Cambridge handbook of situated cognition*, 2009, ch. Problem solving and situated cognition, pp. 264–306.

[27] T. Carraher, D. Carraher, and A. Schliemann, "Mathematics in the streets and in the schools," *British journal of developmental psychology*, vol. 3, pp. 21–29, 1985.

[28] T. Ormerod and L. Ball, *Empirical studies of programmers*, 1993, ch. Does programming knowledge or design strategy determine shifts of focus in prolog programming?, pp. 162–186.

[29] P. Burgess, J. Simons, L. Coates, and S. Channon, *The cognitive psychology of planning*, 2005, ch. The search for specific planning processes, pp. 199–227.

[30] K. Christoffersen and D. Woods, "How to make automated systems team players," *Advances in Human Performance and Cognitive Engineering Research*, vol. 2, 2002.

[31] D. Braziunas and C. Boutilier, "Elicitation of factored utilities," *AI Magazine*, vol. 29, no. 4, pp. 79–92, 2008.

[32] D. E. Smith, "Planning as an iterative process," in *Proc. of AAAI*, 2012.

[33] D. Höller, G. Behnke, P. Bercher, and S. Biundo, "Assessing the expressivity of planning formalisms through the comparison to formal languages," in *Proc. of ICAPS*, 2016.

[34] G. Behnke, D. Höller, P. Bercher, and S. Biundo, "Change the plan - how hard can that be?" in *Proc. of ICAPS*, 2016.

[35] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 39, no. 2, pp. 230–253, June 1997.

[36] M. Walch, K. Mühl, J. Kraus, T. Stoll, M. Baumann, and M. Weber, *From Car-Driver-Handovers to Cooperative Interfaces: Visions for Driver–Vehicle Interaction in Automated Driving*, Cham, 2017, pp. 273–294.

[37] F. Nothdurft, F. Richter, and W. Minker, "Probabilistic human-computer trust handling," 2014.

[38] B. M. Muir, "Trust in automation: Part i. theoretical issues in the study of trust and human intervention in automated systems," *Ergonomics*, vol. 37, no. 11, pp. 1905–1922, 1994.

[39] B. Lim, A. Dey, and D. Avrahami, "Why and why not explanations improve the intelligibility of context-aware intelligent systems," 2009.

[40] J. Lyons, K. Koltai, N. Ho, W. Johnson, D. Smith, and R. Shively, "Engineering trust in complex automated systems," *Ergonomics in Design: The Quarterly of Human Factors Applications*, vol. 24, no. 1, pp. 13–17, 2016.

[41] B. Seegebarth, F. Müller, B. Schattenberg, and S. Biundo, "Making hybrid plans more clear to human users – a formal approach for generating sound explanations," in *Proc. of ICAPS*, 2012.

[42] M. Göbelbecker, T. Keller, P. Eyerich, M. Brenner, and B. Nebel, "Coming up with good excuses: What to do when no plan can be found," in *Proc. of ICAPS*, 2010.