Assumptions of Decision-Making Models in AGI

Pei Wang[1] and Patrick Hammer[2]

[1] Department of Computer and Information Sciences, Temple University
1925 North 12th Street, Philadelphia, PA 19122, USA
[2] Institute of Optimization and Discrete Mathematics, Graz University of Technology
Steyrergasse 30, 8010 Graz, Austria

Abstract. This paper analyzes the assumptions of the decision making models in the context of artificial general intelligence (AGI). It is argued that the traditional approaches, exemplified by decision theory and reinforcement learning, are inappropriate for AGI, because their fundamental assumptions on available knowledge and resource cannot be satisfied here. The decision making process in the AGI system NARS is introduced and compared with the traditional approaches. It is concluded that realistic decision-making models must acknowledge the insufficiency of knowledge and resources, and make assumptions accordingly.

1 Formalizing decision-making

An AGI system needs to make decisions from time to time. To achieve its goals, the system must execute certain operations, which are chosen from all possible operations, according to the system's beliefs on the relations between the operations and the goals, as well as their applicability to the current situation.

On this topic, the dominating normative model is decision theory [12, 3]. According to this model, "decision making" means to choose one action from a finite set of actions that is applicable at the current state. Each action leads to some consequent states according to a probability distribution, and each consequent state is associated with a utility value. The rational choice is the action that has the *maximum expected utility* (MEU).

When the decision extends from single actions to action sequences, it is often formalized as a *Markov decision process* (MDP), where the utility function is replaced by a reward value at each state, and the optimal *policy*, as a collection of decisions, is the one that achieves the *maximum expected total reward* (usually with a discount for future rewards) in the process. In AI, the best-known approach toward solving this problem is *reinforcement learning* [4, 16], which uses various algorithms to approach the optimal policy.

Decision theory and reinforcement learning have been widely considered as setting the theoretical foundation of AI research [11], and the recent progress in deep learning [9] is increasing the popularity of these models. In the current AGI research, an influential model in this tradition is AIXI [2], in which reinforcement learning is combined with Solomonoff induction [15] to provide the probability values according to algorithmic complexity of the hypotheses used in prediction.

Every formal model is based on some fundamental assumptions to encapsulate certain beliefs about the process to be modeled, so as to provide a coherent foundation for the conclusions derived in the model, and also to set restrictions on the situations where the model can be legally applied. In the following, four major assumptions of the above models are summarized.

The assumption on task: The task of "decision making" is to select the best action from all applicable actions at each state of the process.

The assumption on belief: The selection is based on the system's beliefs about the actions, represented as probability distributions among their consequent states.

The assumption on desire: The selection is guided by the system's desires measured by a (utility or reward) value function defined on states, and the best action is the one that with the maximum expectation.

The assumption on budget: The system can afford the computational resources demanded by the selection algorithm.

There are many situations where the above assumptions can be reasonably accepted, and the corresponding models have been successfully applied [11, 9]. However, there are reasons to argue that artificial general intelligence (AGI) is not such a field, and there are non-trivial issues on each of the four assumptions.

Issues on task: For a general-purpose system, it is unrealistic to assume that at any state all the applicable actions are explicitly listed. Actually, in human decision making the evaluation-choice step is often far less significant than diagnosis or design [8]. Though in principle it is reasonable to assume the system's actions are recursively composed of a set of basic operations, decision makings often do not happen at the level of basic operations, but at the level of composed actions, where there are usually infinite possibilities. So decision making is often not about *selection*, but *selective composition.*

Issues on belief: For a given action, the system's beliefs about its possible consequences are not necessarily specified as a probability distribution among following states. Actions often have unanticipated consequences, and even the beliefs about the anticipated consequences usually do not fully specify a "state" of the environment or the system itself. Furthermore, the system's beliefs about the consequences may be implicitly inconsistent, so does not correspond to a probability distribution.

Issues on desire: Since an AGI system typically has multiple goals with conflicting demands, usually no uniform value function can evaluate all actions with respect to all goals within limited time. Furthermore, the goals in an AGI system change over time, and it is unrealistic to expect such a function to be defined on all future states. How desirable a situation is should be taken as part of the problem to be solved, rather than as a given.

Issues on budget: An AGI is often expected to handle unanticipated problems in *real time* with various time requirements. In such a situation, even if the decision-making algorithms are considered as of "tractable" computational complexity, they may still fail to satisfy the requirement on response time in the given situation.

None of the above issues is completely unknown, and various attempts have been proposed to extend the traditional models [13, 22, 1], though none of them has rejected the four assumptions altogether. Instead, a typical attitude is to take decision theory and reinforcement learning as *idealized models* for the actual AGI systems to approximate, as well as to be evaluated accordingly [6].

What this paper explores is the possibility of establishing normative models of decision making without accepting any of the above four assumptions. In the following, such a model is introduced, then compared with the classical models.

2 Decision making in NARS

The decision-making model to be introduced comes from the NARS project [17, 18, 20]. The objective of this project is to build an AGI in the framework of a reasoning system. Decision making is an important function of the system, though it is not carried out by a separate algorithm or module, but tightly interwoven with other functions, such as reasoning and learning. Limited by the paper length, the following description only briefly covers the aspects of NARS that are directly related to the current discussion.

NARS is designed according to the theory that "intelligence" is the ability for a system to be *adaptive* while working with *insufficient knowledge and resources*, that is, the system must depend on *finite* processing capability, make *real-time* responses, *open* to unanticipated problems and events, and *learn* from its experience. Under this condition, it is impossible for the truth-value of beliefs of the system to be defined either in the model-theoretic style as *the extent of agreement with the state of affairs*, or in the proof-theoretic style as *the extent of agreement with the given axioms*. Instead, it is defined as *the extent of agreement with the available evidence* collected from the system's experience.

Formally, for a given statement S, the amount of its *positive evidence* and *negative evidence* are defined in an idealized situation and measured by amounts w^+ and w^-, respectively, and the total amount evidence is $w = w^+ + w^-$. The truth-value of S is a pair of real numbers, $\langle f, c \rangle$, where f, *frequency*, is w^+/w so in [0, 1], and c, *confidence*, is $w/(w + 1)$ so in (0, 1). Therefore a belief has a form of "$S\langle f, c \rangle$". As the content of belief, statement S is a sentence in a formal language *Narsese*. Each statement expresses a relation among a few concepts. For the current discussion, it is enough to know that a statement may have various internal structures for different types of conceptual relation, and can contain other statements as components. In particular, *implication* statement $P \Rightarrow Q$ and *equivalence* statement $P \Leftrightarrow Q$ express "If P then Q" and "P if and only if Q", respectively, where P and Q are statements themselves.

As a reasoning system, NARS can carry out three types of inference tasks:

Judgment. A judgment also has the form of "$S\langle f, c \rangle$", and represents a piece of new experience to be absorbed into the system's beliefs. Besides adding it into memory, the system may also use it to revise or update the previous beliefs on statement S, as well as to derive new conclusions using various inference rules (including *deduction, induction, abduction, analogy*, etc.). Each

rule uses a truth-value function to calculate the truth-value of the conclusion according to the evidence provided by the premises. For example, the *deduction rule* can take $P\langle f_1, c_1\rangle$ and $P \Rightarrow Q \ \langle f_2, c_2\rangle$ to derive $Q\langle f, c\rangle$, where $\langle f, c\rangle$ is calculated from $\langle f_1, c_1\rangle$ and $\langle f_2, c_2\rangle$ by the truth-value function for deduction.[3] There is also a *revision rule* that merges distinct bodies of evidence on the same statement to produce more confident judgments.

Question. A question has the form of "S?", and represents a request for the system to find the truth-value of S according to its current beliefs. A question may contain variables to be instantiated. Besides looking in the memory for a matching belief, the system may also use the inference rules *backwards* to generate derived questions, whose answers will lead to answers of the original question. For example, from question Q? and belief $P \Rightarrow Q \ \langle f, c\rangle$, a new question P? can be proposed by the *deduction rule*. When there are multiple candidate answers, a *choice rule* is used to find the best answer among them, based on truth-value, simplicity, and so on.

Goal. A goal has the form of "S!". Similar to logic programming [5], in NARS certain concepts are given a *procedural interpretation*, so a goal is taken as a statement to be achieved, and an operation as a statement that can be achieved by an executable routine. The processing of a goal also includes backward inference guided by beliefs that generates derived goals. For example, from goal Q! and belief $P \Rightarrow Q \ \langle f, c\rangle$, a new goal P! can be proposed by the *deduction rule*. If the content of a goal corresponds to an executable *operation*, the associated routine is invoked to directly realize the goal, like what a Prolog *built-in predicate* does.

Under the restriction of the available knowledge and resources, no task can be accomplished *perfectly*. Instead, what the system attempts is to accomplish them as much as allowed by its available knowledge and resources. In NARS, decision making is most directly related to the processing of goals, though the other inference activities are also relevant.[4]

In Narsese, an operation is expressed by an *operator* (which identifies the associated routine) with an *argument list* (which includes both input and output arguments). The belief about the execution condition and consequence of an operation is typically represented as "$(condition, operation) \Rightarrow consequence$", which is logically equivalent to "$condition \Rightarrow (operation \Rightarrow consequence)$".[5] This belief can be used in different ways. In an idealized situation (where the uncertainty of the belief and the existence of other beliefs and tasks are ignored), if "$condition$" is true, the execution of "$operation$" will make "$consequence$" true by forward inference; when "$consequence$!" is a goal, backward inference will

---

[3] Since P and Q can be events with an occurence time, the same rules can be used for temporal reasoning, which is described in more detail in [21].

[4] Different types of inference tasks may work together. For example, from important judgments of low confidence, questions can be derived, and from certain questions, goals can be derived, which if pursued give rise to curious and exploratory behaviors.

[5] Like other beliefs, there is a truth-value attached, which is omitted here to simplify the discussion.

generate "*condition*!" as a derived goal. When the latter goal is satisfied (either confirmed by a belief or achieved recursively by other operations), "*operation*!" becomes another derived goal, which is directly achieved by invoking the associated routine. Here the process looks similar to logic programming, though the situation is more complicated, especially in backward inference.

As an open system working in real time, new tasks can come while the system is still working on other goals, and there is no guarantee that all the co-existing goals are consistent with each other in what they demand the system to do. Even if all the innate and given goals are consistent, the derived ones may not be, since they usually come as means to achieve certain goal in isolation, without considering their impacts on the other goals. Even among goals that are consistent with each other in content, they still compete for resources, especially processing time. In NARS, to fully process a goal means to take all relevant beliefs into consideration. Since the system's capability is finite and the goals all should be accomplished as soon as possible, it is usually impossible to fully process all of them. Consequently, it becomes necessary to have *preference* among goals to indicate their different significance to the system.

Instead of defining a separate measurement for preference, NARS takes the "desire as belief" approach [10]. The *desire-value* of statement S is taken as the *truth-value* of statement $S \Rightarrow D$, where D is a *virtual statement* representing the "desired state" where all the system's goals are satisfied. D is "virtual" in the sense that its content is not explicitly spelled out, nor is it actually stored in the system's memory. It is only used in the conceptual design to turn the processing of the desire-values into that of the related truth-values. While every judgement has an assigned truth-value, every goal has an assigned desire-value. Like a truth-value, the intuitive meaning of a desire-value can also be explained using idealized situations. S has desire-value $\langle w^+/w, w/(w+1) \rangle$ if the system believes that if S is realized, w^+ of its w consequences are "good", while the rest of them are "bad", with respect to the system's goals. In this way, the system can calculate the desire-value of a statement according to the desire-value of another statement and the belief that linked them, using the truth-value functions of the inference rules. For example, the desire-value of statement S_1, d_1, is interpreted as the truth-value of statement $S_1 \Rightarrow D$, so can be used with the truth-value of belief $S_2 \Rightarrow S_1$, t_1, by the *deduction* function to calculate d_2, the truth-value of $S_2 \Rightarrow D$, which is the desire-value of statement S_2. In this process the exact content of D is irrelevant, as far as it is the same in its two usages. Even without going into the details of the above calculation, it is easy to see that d_2 depends on both d_1 and t_1. S_2 is highly desired only when S_1 is highly desired and the implication relation $S_2 \Rightarrow S_1$ is strongly supported by available evidence.

Similarly, the *revision rule* can be used to merge conflicting desire-values. For example, after a high desire-value of S_2 is established by a goal S_1, another goal S_3 is taken into consideration, but the system believes that it can be realized only when S_2 is not realized. By deduction again S_2 will get another desire-value d'_2 whose frequency value is low. Now the *revision rule* can combine d_2 and d'_2 into d''_2, as the desire-value of S_2 when both goals S_1 and S_3 are taken into

account. In this case, whether S_2 will still be treated as a goal depends on the total evidence – if the frequency factor in d_2'' is too low, it will not be pursued by the system, despite of the positive evidence from S_1. In this way, "decision making" in NARS can be discussed in two senses:

– In a *narrow sense*, it corresponds to the decision of whether to turn a state-ment 'S' into a goal 'S!'. As explained above, this decision mainly depends on the current desire-value of S, especially the *frequency* factor in it.
– In a *broad sense*, it corresponds to the process in which the related factors, especially the desire value of S, are decided. As explained above, this process may consist of many steps and involve many tasks and beliefs.

As argued in several other publications (such as [18, 20]), since NARS takes a *unified approach* toward AGI, many cognitive functions are interwoven in it and carried out by the same underlying process, rather than by interconnected modules. This is true for functions like "reasoning", "learning", as well as "de-cision making". What is special about the latter is that the decision of whether to pursue a new goal is a binary commitment, a "to do or not to do" matter, which is based on the current beliefs and desires that are all matters of degree.

3 Comparison and discussion

The objective of this paper is not to explain the details of decision-making in NARS, but to use NARS as an example to show the feasibility of building a normative model of decision-making where none of the four traditional assump-tions is made. In the following each of these assumptions is contrasted with the corresponding assumption in NARS.

The assumption on task

Decision making eventually is about the selection of an action among alterna-tives. However, it does not have to be directly formalized as a selection algorithm with a list of alternatives as input – finding the alternatives is also important.

NARS takes the decision to be "Whether to pursue a goal", so it also covers the situation where a goal is not directly achieved by an operation, but via derived goals. Here "to execute an operation" is just a special case of "to achieve a goal", where the content of the goal happens to be an operation. On the contrary, the traditional models leave the derivation of goals out of the model, even though it is a crucial aspect of decision making.

For a given goal, NARS considers each alternative one by one, and each of which is a derived goal, with operation as a special case. In each case, the comparison is between "pursue the goal" and "not to pursue the goal". For a given goal, the system may pursue zero, one, or multiple derived goals, and some of the alternatives may be discovered or constructed in the process. Unlike in the traditional models, in this approach there is no demand for an exhaustive list of mutually exclusive actions to be available in advance for each decision.

The traditional decision-making process can still be carried out in NARS as a special case. If all possible actions are listed, and only one of them can be selected, then the evidence favoring one action will be taken as evidence against the other actions. Consequently, the best action selected by the traditional model will also be the one selected by the *choice rule* of NARS, and its selection will block the others under the mutual-exclusion restriction.

The assumption on belief

In all the models the selection of action is based on the system's relevant beliefs about its *preconditions* and its *effects*. In the traditional models, these two aspects are embedded in the *states* of the environment, while in NARS they are expressed by *statements*. In general, a statement only partially specifies a state.

Based on the assumption of insufficient knowledge, in NARS even if a belief $(condition, operation) \Rightarrow consequence$ has a relatively high frequency and confidence, *condition* does not necessarily specifies the operation's full preconditions, nor *consequence* its full effects. This approach is taken, not because the "state-based" approach is bad, but because it is unrealistic. Even POMDP (partially observable Markov decision process) models are too idealized on this aspect, where states still need to be estimated from observations, since the Markov property is defined only in a state-based representation. There have been attempts in reinforcement learning study to change the "flat" state space into a hierarchical one. However, the current approaches all assume *static* abstractions, and how to get *dynamic* abstractions is still acknowledged as an open problem [1]. For a general purpose system, it is crucial to move between different levels of abstractions, as well as to generate them at run time. A statement-based description satisfies such a need. An AGI should be able to work in a non-stationary environment, where the states of the environment never accurately repeat. In such a situation, though it still makes sense to talk about "the state of the environment", to use them to specify an operation is not possible, because future states are usually different from past ones. A statement, on the other hand, only captures certain aspect of states, so can be repeatedly observed in experience. If a classifier is used to merge similar states, then it actually turns the model into "statement-based", since here one "state" may correspond to different situations.

Another difference between NARS and the traditional models is that the *truth-value* in NARS is not *probability*. This topic has been discussed in previous publications [18, 20], so only the major arguments are summarized:

– In NARS, truth-value measures degree of belief according to available evidence, which may change over time without converging to a limit, so is not a function defined on the space of statements.
– Each statement is evaluated according to its own available evidence, so all the truth-values of the beliefs do not necessarily form a consistent probability distribution, and there can be inconsistency among beliefs.
– Because of the above reasons, the calculation and revision of truth-value cannot be handled as Bayesian conditioning or other theorems of probability

theory. Instead, their calculations are based on the semantics of NARS, using extended Boolean functions.

Once again, this is the case, because under the assumption of insufficient knowledge and resources, an AGI usually cannot maintain a consistent probability distribution among its beliefs, no matter how preferred such a situation is.

In particular, Solomonoff probability [15, 2] is not accepted in NARS, since it demands unrealistic resources, as well as is justified by interpreting "Occam's Razor" as "Simple hypotheses are preferred because they are more likely to be true". In NARS, "Occam's Razor" is interpreted as "Simple hypotheses are preferred because they demands less resources". Therefore, it is respected in resource allocation, but not in truth-value calculation.

The assumption on desire

As a derivative of *belief*, in NARS *desire* is also defined on *statements*, rather than on *states*. Beside the partiality nature, it has additional desirable features.

In both decision theory and reinforcement learning, it is assumed that the states are evaluated by a single *value* function (utility function or reward function). Though such a treatment is natural for many practical problems, it has trouble to be extended into general-purpose systems where there are usually multiple goals at the same time. Because of the conflicts and competitions among the goals, the "value" of an action usually should be judged according to its effects on all the existing goals. When these goals are incommensurable, it is hard to justify such a function. The traditional models simply assume the existence of the function, so as to put the burden on the practitioners. Introducing a monetary measurement like commonly used is not always natural or justifiable.

NARS does not depend on an overall value function, but defines desire-value on statements, including goals as a subset. When considering each action, in many situations it is hard, if not impossible, to judge its value against the "ultimate goal" or "super goal" of the system. Instead, it is usually judged with respect to each goal that happens to be considered at the time, and more considerations will bring other related goals into account to revise its desire-value, as explained previously. In the way, the overall desire-value is *constructed by* the system, not *given to* it. Furthermore, this construction is adaptive and context-sensitive, so it is not a fixed function defined on states or statements.

In the discussions about the safety of AI, many people have incorrectly assumed that the key is to give the system a "super goal" that is beneficial to human beings. In reality an intelligent system cannot always decide its actions by checking them against such a super goal, due to knowledge-resource restriction. What really matter are the currently active goals, which may have become inconsistent with the initial goal from which they were derived [19].

The assumption on budget

One of the most noticeable characteristics of NARS is the resources restriction in its theoretical foundation. On the contrary, most normative models regard

the insufficiency of computational time and space as implementation issues and exclude them from theoretical considerations [6]. It has been argued in several of our publications that "optimum decisions made with sufficient resource" and "optimum decisions made with insufficient resource" are different problems that demand different solutions. Many cognitive functions observed in the human mind were evolved to deal with resource allocation, such as *attention*, *forgetting*, and *emotion* [7, 14]. Their impacts on decision making cannot be understood if the resource restriction is omitted.

For NARS, to work in *real time* means that each problem *instance* (not problem *type*) has an associated time requirement that can only be determined at runtime, and the available resources of the system cannot be predetermined, since it depends on the co-existent tasks. To work in such a situation, in NARS there is no "decision making algorithm" with a fixed computational complexity. Instead, a decision-making process is formed at runtime by a large number of inference steps, and the number and order of the steps are determined at the moment by many factors, so that even if the same problem occurs at different moments, its processing and even the result may be different. NARS cannot guarantee to consider all possibilities when a decision is made, but it considers as many as the current resource supply allows. A lot of work has been done in reinforcement learning to improve efficiency, but the results are all in the form of fixed algorithms, which cannot handle variable time pressure. NARS on the other hand is designed to work under the assumption of insufficient resources, with insufficient time as special case.

4   Conclusion

Mainly because of the insufficiency of knowledge and resources, certain basic assumptions of the traditional models of decision making, namely decision theory and reinforcement learning, cannot be accepted in realistic situations where AGI systems should work. These assumptions include:

– "Decision making" means to select the best action from all applicable actions.
– Beliefs on actions are expressed as probabilistic transitions among states.
– Desires are measured by a value function defined on states.
– The system can afford the resources demanded by the involved algorithms.

Though the traditional models can still be extended and revised, they cannot drop all these fundamental assumptions without becoming fundamentally different models.[6] They should not be taken as idealized models to be approximated, since these assumptions change the nature of the problem of decision making.

The practice of NARS shows that it is feasible to put decision making in a normative model that is based on the assumption of insufficient knowledge and resources. Such a model shares many features with the human mind, while is still justified according to certain principles of rationality. This direction is similar

---

[6] There is no space in this paper to discuss approaches where some of them are rejected.

to ideas like Simon's "bounded rationality" [13], except those ideas are rarely formalized and specified in details to be implementable in computers. Compared to them, NARS provides a much more detailed model, which is also implemented and under testing.[7]

References

1. Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning. Discrete Event Dynamic Systems 13, 41–77 (2003)
2. Hutter, M.: Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability. Springer, Berlin (2005)
3. Jeffrey, R.C.: The Logic of Decision. McGraw-Hill, New York (1965)
4. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement learning: a survey. Journal of Artificial Intelligence Research 4, 237–285 (1996)
5. Kowalski, R.: Logic for Problem Solving. North Holland, New York (1979)
6. Legg, S., Hutter, M.: Universal intelligence: a definition of machine intelligence. Minds & Machines 17(4), 391–444 (2007)
7. Medin, D.L., Ross, B.H.: Cognitive Psychology. Harcourt Brace Jovanovich, Fort Worth (1992)
8. Mintzberg, H., Raisinghani, D., Théorêt, A.: The structure of 'unstructured' decision processes. Administrative Sciences Quarterly 21, 246–275 (1976)
9. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D.: Human-level control through deep reinforcement learning. Nature 518(7540), 529–533 (Feb 2015)
10. Price, H.: Defending desire-as-belief. Mind 98, 119–127 (1989)
11. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River, New Jersey, 3rd edn. (2010)
12. Savage, L.J.: The Foundations of Statistics. Wiley, New York (1954)
13. Simon, H.A.: Models of Man: Social and Rational. John Wiley, New York (1957)
14. Simon, H.A.: Motivational and emotional controls of cognition. Psychological Review 74, 29–39 (1967)
15. Solomonoff, R.J.: A formal theory of inductive inference. Part I and II. Information and Control 7(1-2), 1–22,224–254 (1964)
16. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, Massachusetts (1998)
17. Wang, P.: Non-Axiomatic Reasoning System: Exploring the Essence of Intelligence. Ph.D. thesis, Indiana University (1995)
18. Wang, P.: Rigid Flexibility: The Logic of Intelligence. Springer, Dordrecht (2006)
19. Wang, P.: Motivation management in AGI systems. In: Proceedings of the Fifth Conference on Artificial General Intelligence. pp. 352–361 (2012)
20. Wang, P.: Non-Axiomatic Logic: A Model of Intelligent Reasoning. World Scientific, Singapore (2013)
21. Wang, P., Hammer, P.: Issues in temporal and causal inference. In: Proceedings of the Eighth Conference on Artificial General Intelligence (2015)
22. Weirich, P.: Realistic Decision Theory. Oxford University Press, New York (2004)

---

[7] For source code and working examples, visit *https://github.com/opennars/opennars*.