Advances in Cognitive Systems X (2018) 1-20 Submitted X/2018; published X/2018

The Role of Goal Ranking and Mood-Based Utility in Dynamic Replanning Strategies

Bart Gajderowicz BARTG@MIE.UTORONTO.CA
Mark S. Fox MSF@MIE.UTORONTO.CA
Michael Grüninger GRUNINGER@MIE.UTORONTO.CA
Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road,Toronto, ON M5S 3G8, Canada

Abstract

In classical AI planning, replanning strategies are used to re-evaluate a plan during execution. For human-like agents, goal preferences and emotions play an important role in evaluating a plan's progress. However, most existing systems rely on predefined goal ordering and a static association between an event and its emotion-based utility calculation. During execution, utility of individual actions are used to trigger the replanning process. This approach assumes that a complete sequence of actions can be generated, preferences are known, are transitive, the emotion-based utility of every action's outcome is known, and a replanning condition is well defined. This paper presents an alternative approach, one that does not make assumptions about the agent's or observer's omniscience about factors influencing decision making. Our approach recognizes the bounds that limit the agent and observer equally. To accommodate these limits, first, human-centric goal ranking are grounded in a domain-specific mapping to Maslow's hierarchy. Second, a new replanning condition is proposed with dynamically changing mood-based utility during plan execution.

1. Introduction

The objective of our work is to reproduce behaviour of real-life individuals that interact with social service program for the purpose of evaluating such programs. Towards this end, the framework proposed here strives to *emulate* behaviour exhibited by social service clients as captured by data about participants of a social service program. This paper proposes a human-centric replanning condition based on basic human needs and an mood-based evaluation of a plan by a human-like agent. Traditionally during the planning phase, the agent is assumed to be rational, relying on optimal goal preferences for a given task and the neoclassical utility function that increases expectation of success at each time step (Camerer & Lovallo, 1999; Marnet, 2005). However, due to bounded rationality, a true representation of a human-like agent is not able to consider all information required to make a globally optimal choice that maximizes utility (Simon, 1955). Instead, the chosen plan is an approximation of a *potentially* executable plan. During the plan execution phase, the deviation of the executed plan form the chosen plan takes a toll on the agent (Camerer et al., 2004; Kensinger & Schacter, 2008; Marnet, 2005). Depending on the resilience of the agent, a sufficiently pessimistic evaluation of the executed plan triggers the replanning process. In the proposed framework, rather than relying on a static assignment of valence for each action and outcome state *a priori* that impacts mood, valence is assigned dynamically during execution and creates a dynamic mood-based replanning trigger.

This paper extends the BRAMA framework, described elsewhere (Gajderowicz et al., 2017a) and summarized in the next section, that reproduces decisions of cognitively bounded and seemingly "irrational" agents, focusing on the homeless population. The main contributions of this paper are:

- the extension of expected utility function by incorporating Maslow's (1943) hierarchy (MH) for ranking goals,
- extension of a dynamic mood-based utility function for changing moods over time called the Emotional Cycle of Change (ECOC) (Gajderowicz et al., 2017b),
- a replanning condition based on pessimistic and optimistic moods of the agent, and
- a simulation environment that controls plan execution and plan utility during the execution phase.

In section 2 the BRAMA framework is introduced. In section 3 bounded plan generation and human-like goal ranking are introduced. In section 4 expected utility functions used by BRAMA are introduced. Section 5 describes the replanning process. We finish with related work and concluding remarks in sections 6 and 7.

2. Review of the BRAMA Framework

The **B**ounded **R**ational **A**gent **M**otiv**A**tions framework (BRAMA) is used to construct a bounded agent model and their environment to emulate behaviour of human-like agents and simulate the interaction with their environment (Gajderowicz et al., 2017b). Before moving forward, we make a clear distinction between emulating and simulating social service clients. *Emulation* is the reproduction of an individual's plan generation and plan monitoring processes as the individual interacts with their environment over time. *Simulation* of an individual is the replication of their environment's impact on plan execution over time. The objective of our work is to emulate behaviour of homeless clients according to data about their participation in an intervention program. The simulation environment represents the dynamic and constrained social services used by clients and the impact it has on plan utility after plan execution over time.

The core theoretical tenants of this work are based on behaviour theories used in AI, economics, sociology, and psychology to reproduce behaviour of human-like agents (Gajderowicz et al., 2017a). Economists and AI practitioner focus on understanding the internal processes of decision-making, which we call the "reasoning view": an objective understanding of choices (Etzioni, 1988; Russell, 1997). Within psychology and sociology, rationality is a reference point, and researchers focus on interpreting observed behaviour, which we call the "behavioural view": a subjective understanding of choices (Simon, 1996; Etzioni, 1988). The work of Becker and later Shultz identifies this division as being between the subjective rationality of an emotional agent and the objective rationality of an observer (Zafirovski, 2005). The objective of our work is to emulate seemingly "irrational" behaviour using a rational reasoner by combining these two views in the BRAMA framework.

In BRAMA, the two perspectives are split between two phase of behaviour, planning and execution. During the planning phase the agent is assumed to be rational in the neoclassical sense, to know required information, be non-emotional, and plan their actions based on their own preferred ranking of goals. During the execution phase, however, the agent must adjust to the real and unforeseen consequences of their actions. The responses may be expected and included in the original plan or unexpected, causing the executed plan to deviate from the original plan.

The target domains for a BRAMA agent model are those for which the motivations and means of individual agents are not well understood. This paper targets social service clients, specifically the homeless population. To emulate true homeless clients, the agent model configurations are based on a dataset provided by the Calgary Homeless Foundation (CHF) [1] that captured information about clients as they participate in a "Housing First" (HF) intervention program administered by CHF. The CHF-HF dataset contains information on approximately 4,000 unique clients that participated in the HF program in Calgary, Canada from 2009 to 2015. A complete description of the data and analysis is provided in (Gajderowicz et al., 2018b). Participants were surveyed at program intake with follow-up interviews every three months until exiting the program. Amongst other information, the questionnaire captured client demographics and requests for goods and services. The 763 different types of requests provided by clients were combined into 58 different need categories.

To provide a more objective ranking of goals and calculate utility, BRAMA relies on a domain-specific mappings of agent goals to levels of Maslow's (1943) hierarchy grounds the rankings in basic human needs. There is some consensus amongst researchers that behaviour models can rely on theories like Maslow's hierarchy (MH), given appropriate adjustments for specific domains (Kenrick et al., 2010; Sumerlin, 1995; Henwood et al., 2015). MH categorizes goals into five levels of needs that capture short- and long-term needs of an agent. BRAMA utilizes homeless-specific mapping described by Gajderowicz et al. (2018b) to map requests reported in the CHF-HF data to one or more levels of Maslow's hierarchy. The Ontology of Social Service Needs was developed to map each need category to an MH level based on goal definitions and fifteen key client demographics (Gajderowicz et al., 2018a).

Finally, emotions also play an important part in calculating expected utility. Many existing models assume that emotions can be assigned statically to all observed situations, and that such assignments are known *a priori* (Lin et al., 2012; Steunebrink et al., 2007). However, an agent will respond differently to similar situations depending on their emotional *mood*. Such dynamism in an agent's emotional *mood* is not captured by existing models of emotion-based utility functions due to statically assigned emotions for specific events and scenarios. BRAMA relies on the Emotional Cycle of Change (ECOC) (Gajderowicz et al., 2017b) theory. ECOC recognizes that, in response to the same event or scenario, behaviour changes according to stages of pessimism and optimism and generalizes that process. A replanning condition is triggered when the agent's pessimism threshold is reached. Once triggered, goals are reranked and a new plan is generated.

3. Human-Centric AI Planning

As discussed in section 2, BRAMA adopts two perspectives of rationality, the presumed rational view and the observed behavioural view that considers environmental and internal constraints that may not be known. This section introduces how a bounded agent represents goals and actions

1. The Calgary Homeless Foundation: `http://calgaryhomeless.com/`.

for generating a plan during the planning phase and how goal ranking is incorporated into utility calculation.

3.1 Bounded Plan Generation

According to Simon (1955), bounded rationality limits our ability to be completely rational when making decisions. Three bounds limit the creation of a search tree in BRAMA (Gajderowicz et al., 2017a). The knowledge bound $BR(I)$ limits the amount of correct information an agent can store in memory. The time bound $BR(T)$ limits the number of states that can be visited during the planning phase while creating the search tree. Cognitive limitation set by $BR(C)$ limit the depth of a search tree that can be evaluated. $BR(I)$ impacts how goals are stored and reranked in memory during the planning process, and is discussed here in detail.

$$BR(I) = \{S_{BR,t}, G_t, AS_{BR}\} \tag{1}$$

In bounded AI planning, the knowledge bound, as defined in Equation 1, limits the agent's creation of their search tree to only include those states of the world the agent knows about ($S_{BR,t}$) at time step t, an action schema of known actions and their characteristics (AS_{BR}), and what goal states an agent has (G_t) at time step t. As the plan is generated, interim goals are included to satisfy preconditions of actions in the plan. The final search tree is made up of branches representing a set of possible plans (P^x) from which the agent can select.

S_t provides grounded facts that are known to be true at time step t. $S_{BR,t}$ are states the agent believes are true at time step t. G are basic achievement goal states an agent wants to be true and are mapped to an MH level. G^I are interim goal states an agent wants to be true that are not mapped to an MH level but identified as maintenance goals during the planning phase to satisfy preconditions of actions that satisfy existing unsatisfied goals.

$$G_t^U \subseteq G \cup G^I \text{ , where } \varnothing = G_t^U \cap S_{BR,t} \tag{2}$$

$$G_t^S \subseteq G \cup G^I \text{ , where } G_t^S \subseteq S_t \tag{3}$$

$$G_t^{U+S} = G_t^U \cup G_t^S \text{ , where } G \subseteq G_t^{U+S} \tag{4}$$

G_t^U is a set of unsatisfied goal states at time step t, as per Equation 2. The goal set G_t^S is a set of satisfied goal states at time step t, as per Equation 3. The goal set G_t^{U+S} is a set of all goal states for an agent at time step t, as per Equation 4.

An agent's bounded memory to store knowledge in $BR(I)$ limits the number of goals, states, and actions the client can retain at any one time. The actions an agent can use to create a search tree are stored in the bounded action schema AS_{BR}. Each action a in AS_{BR} is defined along with its preconditions (PRE) and postconditions ($POST$). PRE are states that must be true before a is executed while $POST$ are actions that must be true after a is executed. Some action definitions in AS_{BR} are correct, while other definitions may be incorrect. Incorrectly defined actions in AS_{BR} and referred to as incorrect actions with partially defined preconditions and postconditions, as per Equation 6. The set of incorrect actions is AS_{inc}, while the set of correct actions is AS_{cor}. The union of these actions make up the set AS of all possible actions, as defined in Equation 5:

$$AS = AS_{cor} \cup AS_{inc} \tag{5}$$

where $\varnothing = AS_{cor} \cap AS_{inc}$. The function $inc(a)$ transforms a correct action a to an incorrect action a^*, as per Equation 6. Its inverse $inc^-()$ converts an incorrect action to its correct equivalent.

$$a^* = inc(a) \text{ , where } (PRE \neq PRE^*) \text{ or } (POST \neq POST^*) \tag{6}$$

The agent's knowledge about actions then is defined as AS_{BR}, where

$$AS_{BR} \subseteq AS_{cor} \cup AS_{inc} \tag{7}$$

States at time step t that are true before an action is executed are defined as $s_i \in S_t$ while states true after execution at time step $t+1$ are defined as $s_j \in S_{t+1}$. Finally, the agent's reasoner relies on actions available in AS_{BR} to satisfy goals in G^U, given what the client knows about the world in S_{BR}. The final bounded knowledge the agent uses is $BR(I)$, as defined in Equation 1.

3.2 Human-Centric Goal Ranking

Human-centric goal ranking is based on ranking of concrete goals mapped to basic needs identified by Maslow's hierarchy. As discussed in section 2, while this preference may not be used in the planning phase, it is assumed to be true during the plan execution phase. Goal utility is based on two types of preferences, nominal goal ranking and cardinal goals ranking (Wold et al., 1952).

Nominal goal ranking identifies the order goals are preferred in, whether by the agent or based on an MH level. Agent preferred ranking is represented by the $\succ_A$ relation, where $s_i \succ_A s_j$ indicates the agent (A) prefers goal state s_i over s_j. The MH goal ranking relies on MH levels where each level has a different rank. For example, the original hierarchy ranks any goals mapped to the physiological level the highest, while goals mapped to the self-actualization level are ranked the lowest. Hence, MH is captured by the ordering relation $\succ_{mh}$ in Equation 8.

$$physiological \succ_{mh} security \succ_{mh} social \succ_{mh} esteem \succ_{mh} self\text{-}actualization \tag{8}$$

The $rank(pref, s_i)$ function returns the numerical rank for goal state s_i according to its nominal mapping, as per Equation 9.

$$rank(pref, s_i) < rank(pref, s_j) \iff s_i \succ_{pref} s_j \text{ where } pref \in [MH, A] \tag{9}$$

For and agent's preferred ordering, $rank(A, s_i)$ returns the index the agent assigned to the goal state, where $rank(A, s_i) \in \mathbb{Z}$. For Maslow's hierarchy where $pref$=MH, $rank(MH, s_i) \in \{1, 2, 3, 4, 5\}$ for each of the five MH levels. For example, say the goal state for food is mapped to the physiological MH level, where s_i=$food$ and $rank(MH, s_i)$=1. The goal state for having friends is mapped to the social MH level, where s_j=$friends$ and $rank(MH, s_j)$=3. If a goal is mapped to two different levels, it is represented by two separate goal states at different MH levels. For example, receiving "addiction support" generates two separate goal states, say s_m for the physiological level and s_n for the self-actualization level, resulting in $rank(MH, s_m)$=1 and $rank(MH, s_n)$=5. See Gajderowicz et al. (2018a) for a complete discussion on goal mapping. Once all goal states in G are mapped, the appropriate goal ordering can be applied.

Cardinal goal ranking indicates a degree of importance in relation to other goals. The nominal ranking $rank(pref, s_i)$ is used to calculate whether there are preferred goal states that should be satisfied before s_i. For example, assuming that the physiological goal is the most important, and some physiological goal state s_j is still outstanding, any unsatisfied goal state s_i mapped to higher MH-levels should have a lower utility. The degree to which the utility of s_i is lower is relative to its distance from the unsatisfied and higher ranked physiological level of goal s_j, as defined in Equation 11.

$$min(G^U) = rank(pref, s_i), \text{ where for all } s_j \in G^U, rank(pref, s_i) \leq rank(pref, s_j) \tag{10}$$

$$u(pref, s_i) = 1 - \big(rank(pref, s_i) - min(G^U)\big)^{1/e} \tag{11}$$

To calculate cardinal utility for $pref{=}MH$, BRAMA relies on the MH level of goal state s_i in relation to the lowest outstanding MH-level goal. For $pref{=}A$, BRAMA relies on the index of goal state s_i in relation to the index of the lowest goal state outstanding. First, the function $min(G^U)$ defined in Equation 10 returns the minimum $rank(pref, s_i)$ from all outstanding goal states. Second, $u(pref, s_i)$ defined in Equation 11 returns the cardinal utility of the goal state s_i. Here, the difference between $rank(pref, s_i)$ and $min(G_t^U)$ is taken to the power of $1/e$. This reflects logarithmic declining utility of goals at higher levels of the hierarchy, as originally observed by Bernoulli about a declining marginal utility, and adopted by economists like von Neumann (1944), Savage (1954), and others. The resulting $u(pref, s_i)$ function is a cardinal ranking for one goal state relative to another. Suppose s_i and s_j are states in a set ordered by $\succ_{pref}$, then:

$$u(pref, s_i) \geq u(pref, s_j) \iff s_i \succ_{pref} s_j \tag{12}$$

4. Expected Utility Functions for Bounded Agents

In subjective expected utility, an action's utility represents the probability that an action will successfully satisfy the intended goals (Savage, 1954; Jeffrey, 1990). Unless probabilities are known *a priori*, a bounded agent needs to infer them from their perception of success. Two approaches are introduced next, first based on neoclassical expected utility and second on the changing mood of the agent using ECOC.

4.1 Neoclassical Action Expected Utility

As mentioned in section 1, according to neoclassical theories of utility, the rational expectation of success is always increasing, assumed to be improving from one action to another. While different methods capture this type of "improving" pattern, we normalize all neoclassical theories by settling on the "expected value" function in Equation 13. This expected utility is a simple ratio of satisfied goals (G^S) to all goals (G^{U+S}), as depicted in Figure 1.

$$exp() = \frac{|G^S|}{|G^{U+S}|} \tag{13}$$

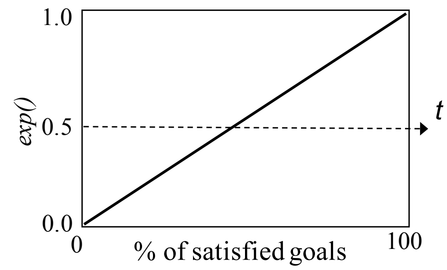


Figure 1: Expected value based on $exp()$.

The final action utility for action a_t^x in plan P^x at time step t is a combination of the postcondition states that satisfy unsatisfied goals, where for each $post_{t,i}^x \in G^U \cap POST_t^x$. Utility of action a_t^x is:

$$u_{exp}(pref, a_t^x) = \frac{\sum_i u(pref, post_{t,i}^x)}{|POST_t^x|} \times exp() \tag{14}$$

4.2 Mood-Based Action Expected Utility

A key limitation of the $exp()$ method and its increasingly monotonic characteristic is that it does not reflect how people's expectations actually change over time (Campbell, 2006; Raiffa, 1961). It turns out that time itself is a factor, where our perception of risk, preferences, reward, or available information changes over time in a way that is not always increasing action utility. Instead, people go through optimistic and pessimistic stages that influences how they perceive expected utility of their choices (Kelley & Connor, 1979; Gajderowicz et al., 2017a). Recall from section 2 that according to the ECOC theory, as individuals begin a task, they are overly optimistic about success, become pessimistic once true efforts becomes apparent, and again become optimistic if sufficient gains towards completing these tasks are made. We can describe these stages in terms of gains and losses of probability of success. During the initial optimistic stages, the probability of success is high without any evidence to justify the optimism. During the pessimism stage, probability of success falls when constraints become apparent. Finally, if constraints are removed, ECOC resembles the $exp()$ function where the probability of success again rises based on new evidence.

The $ecoc()$ utility function in Equation 15 produces the non-monotonic graph in Figure 2, approximating the ECOC graph. $ecoc()$ takes $exp()$ as its only parameter. The result is an adjusted expectation of success according to the ECOC theory.

$$ecoc(x) = \begin{cases} 0.6 - \dfrac{sin(8x-1) + cos(8x)}{x-2}, & \text{if } x \leq 0.8\,; \\ x, & \text{otherwise}\,. \end{cases} \tag{15}$$

The final mood-based utility of action a_t^x takes $exp()$ is a parameter passed to $ecoc()$, multiplied by goal state utility assigned to the action's postconditions:

$$u_{ecoc}(pref, a_t^x) = \frac{\sum_i u(pref, post_{t,i}^x)}{|POST_t^x|} \times ecoc(exp()) \tag{16}$$

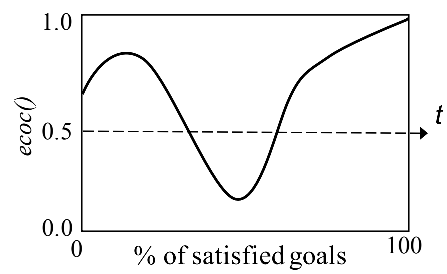


Figure 2: Approximation of the Emotional Cycle of Change function with $ecoc()$.

4.3 Action Utility Relative to the Plan

The calculation of expected utility assigned to an action is based on the probability of successfully achieving a goal that action satisfies. Since, goals closer to the current time step have higher probability of success their weight is relative to the action's position in the plan. BRAMA assigns weights to actions that satisfy basic goal states mapped to an MH level as well as interim actions.

$$dist(s_i^x, s_j^x) = j - i \tag{17}$$

$$aw_k^x = 1 - \frac{dist(s_t^x, s_k^x)}{dist(s_0^x, s_i^x)} \tag{18}$$

For each basic goal state s_i^x, there is a sub-plan $P_t^x(s_i^x)$ of actions required to satisfy s_i^x from the starting state s_t^x, where t=0. To calculate the distance between arbitrary states s_i^x and s_j^x in Plan P^x, $dist(s_i^x, s_j^x)$ returns the number of actions required to transition from state s_i^x to state s_j^x, as per Equation 17. The weight aw_k^x in Equation 18 relies on this distance to calculate the contribution action a_k^x makes to the goals in plan P^x from the initial state s_t^x to when s_i^x is true. A higher weight aw_k^x means action a_k^x is closer to the target goal state s_i^x in sequence x, and contributes more.

4.4 Plan Utility for the Planning Phase with Agent Preferred Ranking

The utility of a plan P^x combines the contribution of each action towards goals being satisfied. During the planning phase, the agent assumes they are objective and rational in the neoclassical sense. It assumes no biases are used to evaluate each plan in the search tree they created. Hence, during the planning phase the agent relies on $exp()$ to calculate plan utility. Also, the agent uses a custom goal preference ordering $\succ_A$ for basic goals mapped to MH needs, relying on $pref(A, s_i)$ for ranking.

When calculating an action's contribution to a plan, it is possible that one action can satisfy goal states at different time steps of a plan. This scenario occurs when alternative actions exist in the same plan, at different time steps. For example, imagine a homeless shelter offers takeaway meals, where the agent receives a hot meal and sandwiches. Both require the precondition the agent is at the shelter, and both satisfy the same physiological need for food. The agent is free to eat the hot meal first and sandwiches later in the day, and vice versa. The action of eating either food can be applied any time after the food is obtained.

(a) Average action utility $U_{exp}(A, a_t^x)$ for action a in plan P^x as goal states in G^U are being satisfied.

(b) Average plan utility $U_{exp}(A, P_t^x)$ as goal states in G^U are being satisfied.

Figure 3: Average neoclassical action utility.

This type of scenario is possible when alternative actions have overlapping postconditions and preconditions for some plan P^x. Generally speaking, if an action a_t^x originally at time step t in plan P^x can be executed at other time steps k, the action's utility increases. To capture such cases, $U_{exp}(pref, a_t^x)$ calculates the contribution action a_t^x makes to plan P^x at all applicable time steps k where the action can be executed, as defined in Equation 19, and visualized in Figure 3a.

$$U_{exp}(pref, a_t^x) = u_{exp}(pref, a_t^x) \times \sum_t aw_k^x \tag{19}$$

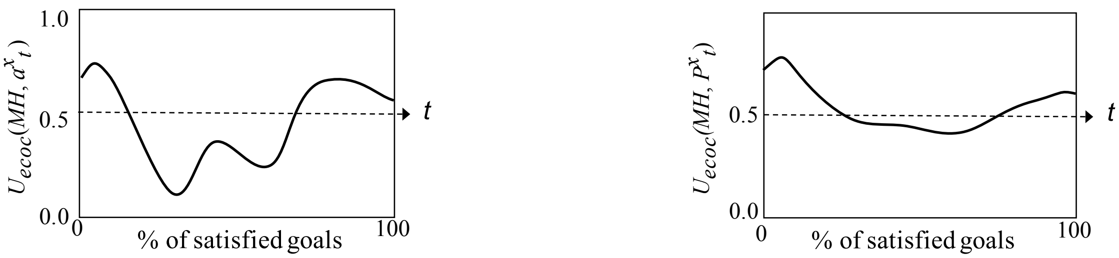
Finally, plan utility combines goal and action utility as the mean utility of all actions in the plan up to time step t, as per Equation 20, and illustrated in Figure 3b. The plan with the highest utility for the entire length of the plan is selected by the agent for execution, as discussed in the next section.

$$U_{exp}(pref, P_t^x) = \frac{\sum_k U_{exp}(pref, a_k^x)}{|P_t^x|} \tag{20}$$

4.5 Plan Utility for the Execution Phase with MH-based Ranking

During the execution phase, realistic consequence of planned actions become apparent. First, the agent's preferred order of basic needs is not used. Rather, MH ranking according to domain-specific mapping to MH levels is used, as defined in the ordering Equation 8. This ensures that regardless what order the agent preferred during the planning phase, reality during execution is represented by Maslow's order. Second, if the agent was too optimistic during the planning phase, they will not improve their utility after each action, as implied by neoclassical theory. Instead, ECOC is used to capture a pessimistic evaluation of their planned actions in response to their true consequences.

If the agent was too optimistic during the planning phase, action utility is calculated using $U_{ecoc}(MH, a_t^x)$, as defined in Equation 21. As illustrated in Figure 21, the trend of action utility is influenced by the ECOC graph in Figure 2 and discussed in section 4.2. The initially high utility represents an optimistic evaluation of actions. As the plan is executed, the evaluation is more



(a) Average action utility $U_{ecoc}(MH, a^x_t)$ for action a in plan P^x as goal states in G^U are being satisfied.

(b) Average plan utility $U_{ecoc}(MH, P^x_t)$ as goal states in G^U are being satisfied.

Figure 4: Average ECOC action utility.

pessimistic resulting in a lower utility. Over time, the plan's utility begins to rise again.

$$U_{ecoc}(MH, a^x_t) = u_{ecoc}(MH, a^x_t) \times \sum_k aw^x_k \tag{21}$$

Finally, plan utility combines ECOC-based goal and action utility as the mean utility of all actions in the plan up to time step t, as per Equation 22, and illustrated in Figure 4b. Unlike the planning phase, plan utility is not used to select the best plan. Rather it is used to indicate when to trigger the replanning process, as discussed in the next section.

$$U_{ecoc}(pref, P^x_t) = \frac{\sum_k U_{ecoc}(pref, a^x_k)}{|P^x_t|} \tag{22}$$

5. Human-Centric Replanning Phase

An important part of emulating a bounded agent's plan generation and monitoring is the environment that forces the agent to reevaluate their plan during plan execution. This section introduces the BRAMA simulation environment that executes an agent's plan and calculates the "true" utility during execution. As the agent monitors the "true" plan utility, a significant difference between the planned and "true" utilities triggers the replanning process. BRAMA incorporates a discrete-event simulation (DES), a popular simulation architecture for bounded agents in social sciences (Harpring et al., 2014). In a DES, the execution of complex systems is represented as an ordered sequence of events. Using forward-chaining, an action is executed as a discrete event at time step t rather than continuously over time. The actual time taken between each event may vary in length.

The simulation environment controls when plan generation, monitoring and execution occur with several modules that perform specific functions, as listed in Table 1. $simulate(S_t, G_t)$ begins the simulation process. It takes two parameters, the current state of the world S_t and the agent's unsatisfied goals G^U_t at time step t. At the start of the simulation, $S_{BR,t}$ and G_t are used to initialize the process. It returns a set of final plans that were executed and resulting world states, PL_{Final} and S_{Final} respectively. $plan(S_t, G_t, AS_{BR})$ generates a plan P^x_t using BRAMA's STRIPS-based

Table 1: Simulation environment modules.

Module	Descriptions
$simulate(S_t, G_t)$	Begins the simulation process.
$plan(S_t, G_t)$	Generates a plan P_t^x using BRAMA's STRIPS-based planner.
$next\_action(P^x)$	Returns the next action a_t^{*x} in plan P_t^x.
$exec(a_t^x, S_t, G_t)$	Executes action a_t^x, given true states S_t and goals G_t. Returns the new states S_{t+1}.
$reduce(G_x)$	Returns a reduces set of goals G_R, as per Equation 23.

planner (Gajderowicz et al., 2017a). The third parameter AS_{BR} ensures the plan is generated using the agent's bounded action schema in $BR(I)$. $next\_action(P_t^x)$ returns the next action a_t^x in plan P^x to be executed. The action must be a correct action in AS_{cor} to ensure realistic preconditions and postconditions are enforced on the agent during the execution phase. To ensure it is correct the inverse of equation 6 is used, mainly $a_t^x{=}inc^-(a_t^{*x})$ where $a_t^x \in AS_{cor}$ whether a_t^{*x} is in AS_{cor} or not. $exec(a_t^x, S_t, G_t)$ executes the action, given the agent's current unsatisfied goals in G_t and the true current state S_t. During, execution, it transitions the state S_t to the new state S_{t+1}.

The agent monitors plan execution by comparing plan utility to their mood-based threshold $ecoc{-}th$. If the utility is above the threshold execution continues. If the utility falls below the threshold the replanning process is triggered. During the replanning process, $G_R{=}reduce(G_t)$ returns a subset of goal states in G_t that remain after other goals were deferred, as defined in Equation 23, and discussed further in the next section. The subset of original goals in G (excluding interim goals) not yet satisfied are returned in the goal set G_R, and used to generate a new plan. For each goal state $s_i \in G_t \cap G$ and each action $a_t^x \in P_t^x$ that satisfies that goal, where $s_i \in POST_t^x$:

$$G_R = reduce(G_t) : \{g_i \in G_t \cup G | g_i \in POST_t^x \wedge U(a_t^x) \geq act\text{-}th\} \tag{23}$$

5.1 BRAMA Agent Model

The BRAMA agent model $\mathbb{M}$ provides properties used to define an agent that generates and monitors plans while the simulation executes them. The $\mathbb{M}$ structure in Equation 24 represents a particulate type of individual and their characteristics. $demo()$ categorizes the agent as some cohort of a population based on their demographics, such as age, gender, or income. The agent's bounds $(BR(I), BR(C), BR(T))$ indicate their cognitive limitation during the plan generation process. The agent's expected utility function during plan generation is $exp()$. During the execution phase, the utility function ($executil$) can be either $exp()$ or $ecoc()$. Their initial goals at time step $t{=}0$ are G_0. During the planning phase, the agent uses their own goal preferences ($pref{=}A$). During the execution phase the simulation can be configured to use the agent's ($pref{=}A$) or Maslow's ranking ($pref{=}MH$). During the planning phase, the agent always maximizes their utility. During the execution phase, the agent can either maximize their their utility ($planutil{=}planutilswap$) or not ($planutil{=}none$). Finally, to trigger the replanning process, the agent has two thresholds $ecoc\text{-}th$ and $act\text{-}th$, as discussed in the next section.

$$\mathbb{M} = \{demo(), BR(I), BR(C), BR(T), executil, G_0, pref, planutil, ecoc\text{-}th, act\text{-}th\} \tag{24}$$

5.2 BRAMA Simulation Environment

Table 2: BRAMA simulation environment attributes.

Attribute	Descriptions
t	Simulation time unit.
S_t	Correct state of the world at time step t.
G_t	Agent goals at time step t, where $G_t \subseteq G^U$.
P^x	Plan at index x.
P_t^x	Partial plan P^x from start to time step t.
a_t^x	Current action being executed.
G_R	Reduced goals, where $G_R \subseteq G_t \cap G$ and $G_R = reduce(G_t)$.
G_D	Deferred goals, where $G_D = G_t - G_R$.
S_R	State after reduced goals are satisfied.
S_D	State after deferred goals are satisfied.
PL_R	Set of partial plans after reduced goals are satisfied.
PL_D	Set of partial plans after deferred goals are satisfied.
S_{Final}	Final state of the world, returned by $simulate()$.
PL_{Final}	Final set of partial plans, returned by $simulate()$.

Figure 5 presents the simulation flowchart. It relies on the agent model $\mathbb{M}$ properties to control how the agent interacts with their environment. The simulation environment attributes and modules are listed in Tables 1 and 2. If the agent is configured to evaluate plan utility during execution using ECOC then $executil{=}ecoc$ and the simulation may trigger the replanning process.

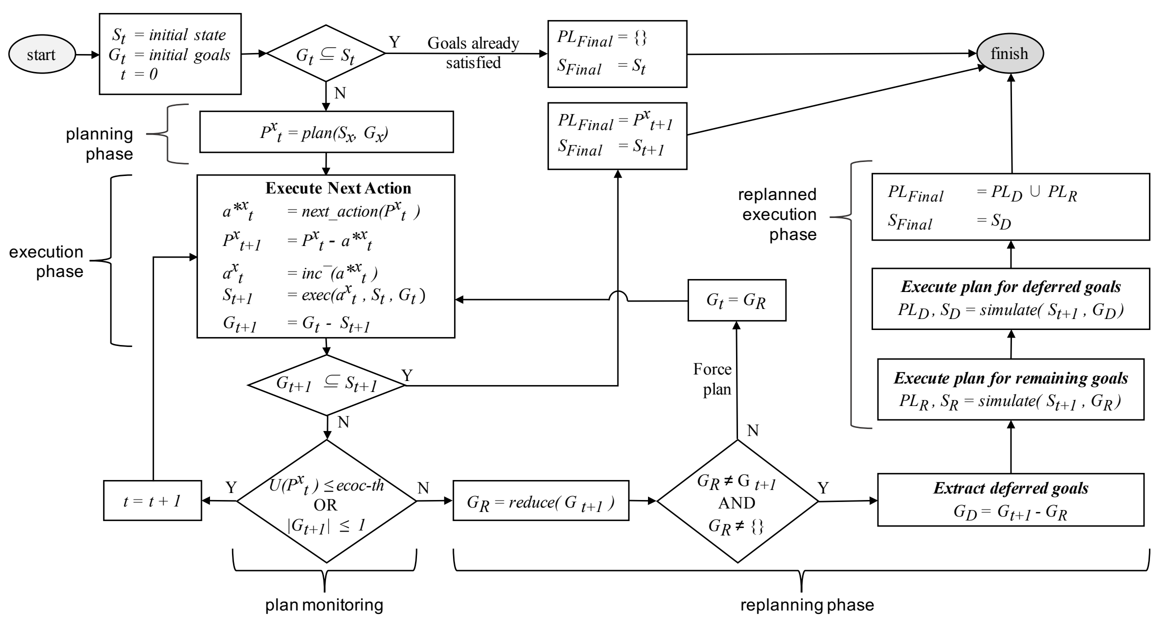


Figure 5: Agent simulation flowchart for $simulate(S_t, G_t)$.

During plan execution, plan utility is recalculated. While monitoring execution, the agent compares the new utility to their $ecoc\text{-}th$ threshold. If $U_{ecoc}(pref, P_t^x) > ecoc\text{-}th$, plan execution continues at time step $t{=}t+1$. Otherwise, goals are reduced to $G_R{=}reduce(G_t)$ according to the agent's $act\text{-}th$ threshold, as per Equation 23. The deferred goals are added to the set G_D. If $U_{ecoc}(pref, P^x) < ecoc\text{-}th$ but either no goals can be removed or all goals are removed, a new forced plan is created and executed without considering $ecoc\text{-}th$. Any goals satisfied by a "forced" plan accumulate up to three times in the final goal count. If a "forced" plan cannot satisfy its goals it becomes a "failed" plan. Goals of a "failed" plan persist until the end of the simulation. After a "forced" plan completes or becomes a "failed" plan, the agent generates a new plan for the remaining goals. This plan's initial state and goals are those that existed the last time step where $U_{ecoc}(pref, P_x) > ecoc\text{-}th$. The cycle continues until all goals are satisfied or the last plan fails.

Once reduced goals are successfully satisfied the state of the world is represented as S_R and plans required to satisfy reduced goals is PL_R. Once deferred goals are retrieved and satisfied, the state of the world is represented as S_D and set of plans required to satisfy all deferred goals is PL_D. The simulation is complete when all reduced and deferred goals are satisfied. The final state of the world is S_{Final}. The set of plans used to satisfy reduced and deferred goals is PL_{Final}, where $PL_{Final}{=}PL_D \cup PL_R$.

5.3 Replanning Example

Figures 6 to 8 illustrate the decision trees an agent creates, and how the replanning process reranks goals. Consider an agent representing a homeless client who's goals are to obtain food (physiological), meet with friends (social) and a housing worker (security). The preferred order if $[friends, housing, food]$. To achieve their goals, the agent creates a search tree with several plans to satisfy them. According to the correct action schema AS_{cor}, food can be obtained by going to the soup kitchen at 11:00 or 12:00. An agent can also purchase food at the store at any time at a cost of \$10.00 and panhandle for more money. The agent can visit friends after the 12:00 lunch at the common area where clients socialize. Finally the agent can book an appointment with a housing worker and wait until they get called in for their appointment, which may take the entire day. The agent's bounded action schema AS_{BR} is a subset of $AS_{cor} \cup AS_{inc}$. In AS_{inc}, the cost of food is \$3.00, which the agent believes rather than the true cost of \$10.00 defined in AS_{cor}.

5.3.1 Step 1: Planning

In Figure 6, the agent begins the planning process in the "Start" state at time step $t{=}0$. With a cognitive bound of $BR(C){=}4$ the agent can look four steps ahead. Each action transitions the agent into a new state at time steps $t{=}1, \ldots, 4$. With a $BR(T){=}24$ the agent can only see the first 24 states in the search tree, omitting the last two branches that begin with a visit to the housing worker and panhandling. After calculating the utility of each plan using $exp()$ and agent's preferred goal ordering for $\succ_A$, the agent chooses plan P^2 as having the highest utility. The order goals are satisfied in remains the preferred order $[friends,\ housing,\ food]$.

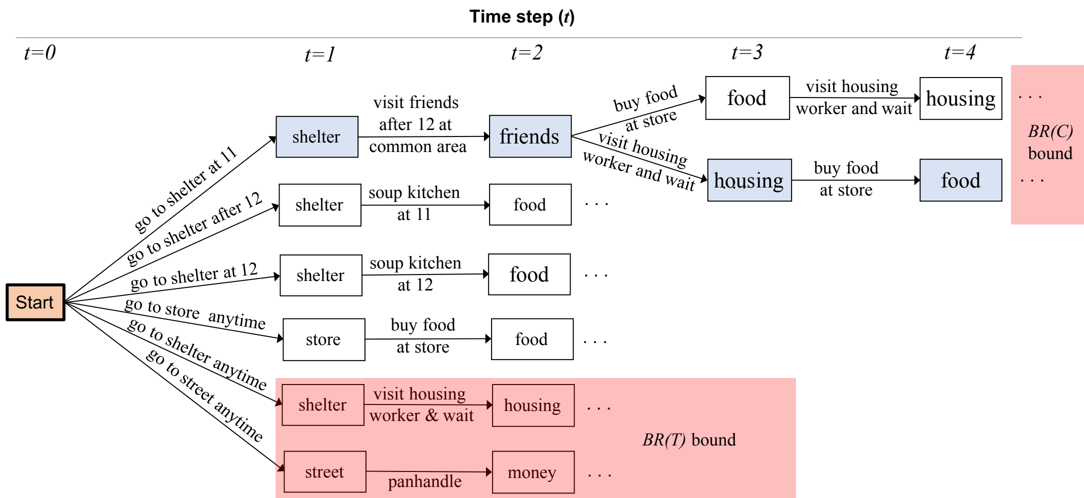


Figure 6: First search tree starting at time step t=0, and selected plan P^2.

5.3.2 *Step 2: Execute Original Plan*

During execution, the agent meets with friends at time step t=2 with the intention of next visiting the housing agent to secure housing, then going to the store to buy food. Once at time step t=2, however, the plan is reevaluated using Maslow's ordering and the $ecoc()$ expected utility function. Since housing (security) is ranked lower than food (physiological) according to MH ranking $\succ_{mh}$ but not according the preferred ranking $\succ_A$, the agent becomes unexpectedly hungry. Knowing they will only have one meal today, they become worried about spending the entire time waiting for the housing worker without a meal beforehand. In this scenario, the utility of P^2 falls below the agent's $ecoc$-th threshold, triggering the replanning process.

5.3.3 *Step 3: Replanning*

The replanning process begins by identifying the actions for which $U_{ecoc}(pref, a_t^x) < act\text{-}th$, and deferring any goals they satisfy. Housing is less important than food in Maslow's order and, say, utility of a "housing" action falls below the threshold, hence $housing$ is deferred. A new search tree is created for the remaining goal, $food$. The agent knows that it is too late for visiting the soup kitchen, as it is after 12:00. There are only two possible plans for which preconditions are true in $S_{BR,2}$ at time step t=2. The first is plan P^{11} where the agent buys a sandwich at the store for \$3.00 with the \$5.00 the agent has. The second is plan P^{12} in which the agent first panhandles for more money then purchases a sandwich at the store. Believing they have enough money for a sandwich and having low expectation of making any money panhandling, the highest utility is calculated for plan P^{11}, as depicted in Figure 7.

Starting at time step t=2, the agent generates and chooses plan P^{11} for execution. Plan execution then begins at time step $t+1$ until replanning is triggered or all goals are satisfied. Taking the union of the partially executed plan $P^2_{0,1}$ and the new plan P^{11} we get $P^{2\cup 11}$ with goal order $[friends, food]$. For the new plan, the already satisfied goal $friends$ remains satisfied. The goal

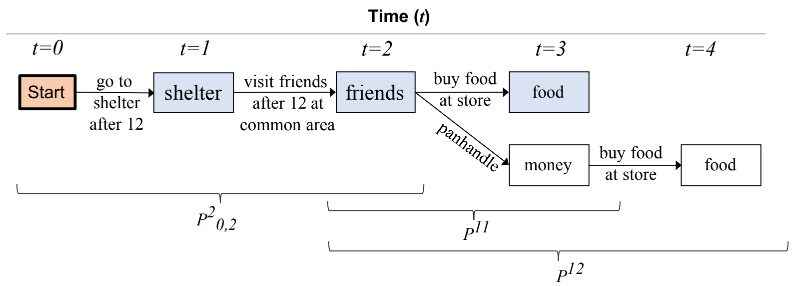


Figure 7: Second search tree after replanning, starting at time step t=2, and selected plan $P^{2 \cup 11}$.

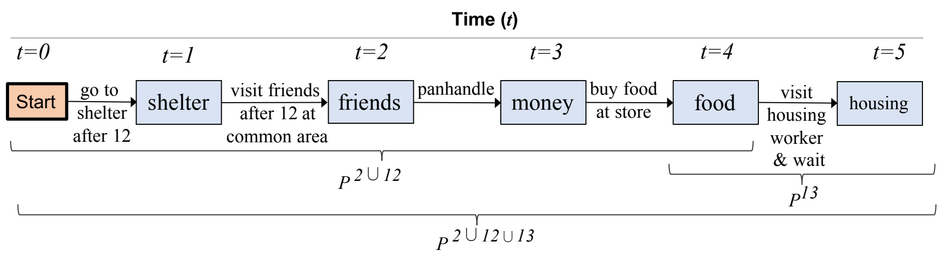


Figure 8: Third search tree after replanning, starting at time step t=4, and selected sub-plan P^{13}, producing the final plan PL_{Final}=$P^{2 \cup 11 \cup 13}$.

$food$ is moved up from the third place to the second place in plan P^2. The goal for $housing$ is deferred, and omitted from $P^{2 \cup 11}$.

5.3.4 Step 4: Execute New Plan

During execution of $P^{2 \cup 11}$, the agent learns the true cost of food at the store is \$10.00 not \$3.00. The agent now tries to overcome this precondition. As described in section 5.1, the agent tries a "forced" plan execution that ignores *ecoc-th*. However the precondition is a hard requirement, and the agent has no immediate actions that allow them to purchase food. Plan $P^{2 \cup 11}$ is deemed a "failed" plan. The agent can try another option up to three times. For example, the agent tries plan P^{12} available to them, to panhandle first then purchase the sandwich, as illustrated in Figure 7. Despite lowered expectations, the agent successfully makes enough money panhandling to buy a sandwich at time step t=4 for \$10.00, concluding execution of the plan $P^{2 \cup 12}$.

5.3.5 Step 5: Planning and Execution for Deferred Goals

Once the $friends$ and $food$ goals have been achieved, deferred goals are retrieved and added to G_t. For this example, the $housing$ goal is retrieved at time step t=4. A new plan is created and executed, adding the new action of visiting the housing worker and waiting for the appointment. Figure 8 illustrates the new plan P^{13} and how it extends the previous plan $P^{2 \cup 12}$. The result is a new plan $P^{2 \cup 12 \cup 13}$ that is ready for execution. The simulation assumes executed actions are success with 100% probability. Hence, after execution at time step t=5 the agent successfully satisfies all goals in G^U. Alternatively, the agent's success could depend on the availability of the housing worker before the day is finished, given some probability, but which is outside of the scope for this paper. The final plan is a union of all partial plans, resulting in plan $P^{2 \cup 12 \cup 13}$, stored in PL_{Final}. The final order of goals is $[friends, food, housing]$ at time step t=5. Through the replanning, the goals were reranked from their original order $[friends, housing, food]$ preferred by the agent.

6. Related Work

Our approach to goal ranking and replanning incorporates ideas from a number of fields required to emulate human-like behaviour. While BRAMA builds on top of this work, existing methods for plan generation, execution, monitoring, and replanning adopt the neoclassical approach to agent behaviour, making it unsuitable for emulation of human behaviour.

The STRIPS-based planner used by BRAMA to generate plans is limited according bounded rationality. Most system incorporate methods for overcoming such bounds and improve efficiency. For example, to overcome time bounds, classic planners like NOAH,NONLIN, SIPE, PRS and PRIAR can execute and monitor multiple operations in parallel, as reviewed by Hendler et al. (1990). A few recent systems explicitly define bounds as part of the planning process. Similar to BRAMA, PUG/X explicitly sets cognitive and time bounds on the search tree, as well as the number of plans to consider before stopping plan search (Langley et al., 2017). Some research into simulating certain types of impairments exist. The Soar system's episodic memory was evaluated to investigate how well it performs when different types of memory loss occur (Nuxoll et al., 2010).

Goal reasoning in BRAMA focuses on ranking and satisfying most important goals first while lower ranked goals are deferred. An agent's preferred goal ranking is used during the planning phase, which are provided *a priori*. A domain-specific mapping of goal to Maslow's hierarchy is used as the basis for goal ranking and utility calculation during the execution phase. There are many systems that rely on a hierarchical goal structure to reason about goals that BRAMA can benefit from, provided a human-centric representation was used. Meneguzzi et al. (2013) propose a hierarchical representation of goals expressed as commitments. ICARUS uses hierarchical goal definitions and a reactive goal management process (Langley & Choi, 2006). In later versions goals can be reprioritized through nominating goals to the top of the hierarchy in certain scenarios (Choi, 2010). Shivashankar et al. (2013) proposed the Hierarchical Goal Network (HGN), a set of predefined methods that define the relations between operators, goals, preconditions, and subgoals. ActorSim is a simulator goal refinement and planner that uses hierarchical goal and task networks (Roberts et al., 2016). Relying on machine learning, the agent learns to perform increasingly sophisticated tasks more efficiently.

Replanning in BRAMA is based on triggering the replanning process, reranking goals, deferring lower ranked goals, and constructing a new plan for the new goal ranking. By concentrating on subset of most important goals, the type of planning performed by BRAMA planner is referred to as partial satisfaction planning (Benton et al., 2009). This approach can be contrasted with systems that retain goals but modifying or repair existing plans. This approach is used in the case-based planning literature. For example, Lee et al. (2008) created a hierarchical case-based reasoner that selects parts of previously used plans to modify the existing plan. HGN uses a similar approach by repairing plans with predefined methods (Shivashankar & Alford, 2013). The PRODIGY system was extended with a replanning module called RAP (Rizzo et al., 1999). RAP depends on personality-driven goal preferences that modify a previously generated plan.

Other systems are similar to BRAMA and rely on goal reasoning to control plan regeneration rather than refinement. Cushing et al. (2008) provide a framework that satisfies abstract goals defined as commitments (required objectives) and opportunities (optional objectives). The replanning stage selects objectives based on which goals should be satisfied to maximize utility and minimize cost. PUG/X triggers replanning when one of four anomaly types are detected during the execution phase that were not found during the planning phase (Langley et al., 2017). Once detected, goals are reranked and a new plan is generated given the current state of the agent. Some systems like PrefPlan (Brafman & Chernyavsky, 2005) and AltAlt^{PS} (van den Briel et al., 2004) rely on predefined common-sense rules that decide when a plan needs to be modified and goals are reprioritized.

Finally, BRAMA relies on a dynamic assignment of utility to goals and actions based on the "mood" of the agent. Similarly, Ojha et al. (2017) propose a replicable, domain-independent computational model for the emotional appraisal of a plan that generalizes the assignment of emotions to events. Such generalization, however, is in contrast to most emotion-based planners, relying instead on predefined associations between emotional appraisals of specific events (Lin et al., 2012). These include AI planners like ACRES/WILL, ActAddAct and EM-ONE. Often, appraisal theory with OCC is used to associate events with discrete emotional responses and valence, including FAtiMA, EM, FLAME, Émile, and work by Gmytrasiewicz et al. (Lin et al., 2012). AI planners like EMA utilize arousal theory and weighted drives to rate the utility of plans. Emotions are also used as replanning triggers. Steunebrink et al. (2007) propose a representation of emotional responses that trigger replanning using predefined deliberation effects based on OCC. While emotions and responses are statically defined, the symbolic representations capture complex combinations of emotions that trigger replanning in a human-like agent.

7. Conclusion and Future Work

The work presented here is based on the theoretical tenets that, for human-like goal-driven agents, planning and execution stages can diverge and trigger a replanning process. It extends earlier work with complete definitions of expected utility functions for planning and execution phases. This paper defines the utility calculation for goal ranking based on a domain-specific mapping of agent goals to Maslow's hierarchy as well as a dynamic mood-based utility calculation. The divergence between the planning and execution utility functions leads to a novel trigger condition for replanning. The result is a set of partial plans that aim to reproduce the changes of seemingly "irrational" agents using a rational reasoner.

However, the BRAMA agent model lacks support for a number of human-like characteristics. First, the planning process is sequential, generating plans one goal at a time. The BRAMA planner would benefit from implementing a parallel planning process to pursue multiple goals at once. Second, the mapping of goals to Maslow's hierarchy levels could be extended by creating a hierarchical goal network, allowing the ranking of goals at different levels of abstraction. Third, the BRAMA agent model would benefit from the ability to ranks existing actions and learn about new actions.

The current agent simply iterates through available actions until one that satisfies outstanding goals is found. By providing actions with weights and probabilities or ability to modify them during execution, the agent could rerank actions, abandon incorrect or unused actions, and discover new actions through exploration. Finally, the agent does not have the ability to generate new goals. Basic goals are provided *a priori* and reranked dynamically. Interim goals are added as required by the planner. The generation of basic goals would allow the agent to find ways of satisfying basic MH needs with new goal states. For example, to satisfy the physiological need for food, the agent could learn of new goal states that satisfy the same underlying physiological needs.

Finally, the BRAMA model would benefit from additional datasets that track the same or different intervention programs. By relying solely on the CHF-HF dataset, the BRAMA model was evaluated on its ability to emulate CHF-HF participants. By calibrating BRAMA models with additional data, a program implementation could be evaluated and ranked against others using the BRAMA framework. For example, in addition to participant demographics and whether they had successful or failed outcomes, program participants across two cities could be categorized by BRAMA agent model characteristics in $\mathbb{M}$ and grouped into appropriate cohorts. By identifying the same $\mathbb{M}$-based cohorts that failed in one city but succeeded in the other, city-specific factors could be identified that lead to the differences in outcomes.

Acknowledgements

This research was funded by NSERC Discovery Grant. The authors would like to thank the reviewers for their insightful comments, as well as the Calgary Homeless Foundation for sharing their dataset with us, and state the our findings do not reflect the views of the Foundation.

References

Benton, J., Do, M. B., & Kambhampati, S. (2009). Anytime heuristic search for partial satisfaction planning. *Artificial Intelligence*, *173*, 562–592.

Brafman, R. I., & Chernyavsky, Y. (2005). Planning with Goal Preferences and Constraints. *Proceedings of the International Conference on Automated Planning and Scheduling* (pp. 182–191).

van den Briel, M., Sanchez, R., & Kambhampati, S. (2004). Over-subscription in Planning: A Partial Satisfaction Problem. *ICAPS 2004 Workshop on Integrating Planning into Scheduling*.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games. *The Quarterly Journal of Economics*, *119*, 861–898.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and Excess Entry : An Experimental Approach. *The American Economic Review*, *89*, 306–318.

Campbell, S. (2006). Risk and the subjectivity of preference. *Journal of Risk Research*, *9*, 225–242.

Choi, D. (2010). Nomination and Prioritization of Goals in a Cognitive Architecture. *Proceedings of the 10th International Conference on Cognitive Modeling* (p. 25).

Cushing, W., Benton, J., & Kambhampati, S. (2008). *Replanning as a Deliberative Re-selection of Objectives*. Technical report, Arizona State University, CSE Department.

Etzioni, A. (1988). Normative-affective factors: Toward a new decision-making model. *Journal of Economic Psychology*, *9*, 125–150.

Gajderowicz, B., Fox, M. S., & Grüninger, M. (2017a). General Model of Human Motivation and Goal Ranking. *2017 AAAI Fall Symposium Series on Standard Model of the Mind* (p. 6). Arlington, VA: AAAI Press.

Gajderowicz, B., Fox, M. S., & Grüninger, M. (2017b). Requirements for Emulating Homeless Client Behaviour. *Proceedings of the AAAI Workshop on Artificial Intelligence for Operations Research and Social Good* (p. 7). San Francisco, CA: AAAI Press.

Gajderowicz, B., Fox, M. S., & Grüninger, M. (2018a). Ontology Of Social Service Needs: Perspective of a cognitive agent. *Proceedings of the 2018 Joint Ontology Workshops, Cognition And OntologieS + Explainable AI* (pp. 1–12). Cape Town.

Gajderowicz, B., Fox, M. S., & Grüninger, M. (2018b). Report on the Ontology of Social Service Needs: Working Paper. From `http://bit.ly/gajderowicz-wp-ossn-2018-pdf`.

Harpring, R., Evans, G. W., Barber, R., & Deck, S. M. (2014). Improving efficiency in social services with discrete event simulation. *Computers & Industrial Engineering*, *70*, 159–167.

Hendler, J. A., Tate, A., & Drummond, M. (1990). AI Planning : Systems and Techniques. *AI Magazine*, *11*, 61–77.

Henwood, B. F., Derejko, K.-S., Couture, J., & Padgett, D. K. (2015). Maslow and Mental Health Recovery: A Comparative Study of Homeless Programs for Adults with Serious Mental Illness. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*, 220–228.

Jeffrey, R. C. (1990). *The logic of decision*. University of Chicago Press.

Kelley, D., & Connor, D. R. (1979). The emotional cycle of change. In *The 1979 annual handbook for group facilitators*. Wiley.

Kenrick, D. T., Griskevicius, V., Neuberg, S. L., & Schaller, M. (2010). Renovating the Pyramid of Needs: Contemporary Extensions Built Upon Ancient Foundations. *Perspectives on psychological science: A Journal of the Association for Psychological Science*, *5*, 292–314.

Kensinger, E. A., & Schacter, D. L. (2008). Memory and Emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions*, chapter 37, 601–617. New York, NY: The Guilford Press, 3rd edition.

Langley, P. W., & Choi, D. (2006). A Unified Cognitive Architecture for Physical Agents. *National Conference on Artificial Intelligence* (pp. 1469–1475).

Langley, P. W., Choi, D., Barley, M., Meadows, B., & Katz, E. (2017). Generating , Executing , and Monitoring Plans with Goal-Based Utilities in Continuous Domains. *Advances in Cognitive Systems* (pp. 1–12).

Lee, C.-h. L., Cheng, K. Y.-r., & Liu, A. (2008). A Case-Based Planning Approach for Agent-Based Service-Oriented Systems. *EEE International Conference on Systems, Man and Cybernetics, 2008. SMC 2008* (pp. 625–630). IEEE.

Lin, J., Spraragen, M., & Zyda, M. (2012). Computational Models of Emotion and Cognition. *Advances in Cognitive Systems*, *2*, 59–76.

Marnet, O. (2005). Behaviour and Rationality in Corporate Governance. *Journal of Economic Issues*, *1*, 4–22.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396.

Meneguzzi, F., Telang, P. R., & Singh, M. P. (2013). A first-order formalization of commitments and goals for planning. *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence* (pp. 697–703).

Nuxoll, A., Tecuci, D., Ho, W. C., & Ningxuan, W. (2010). Comparing Forgetting Algorithms for Artificial Episodic Memory Systems. *Remembering Who We Are - Human Memory for Artificial Agents Symposium*, (pp. 14–20).

Ojha, S., & Williams, M.-A. (2017). Emotional Appraisal: A Computational Perspective. *Advances in Cognitive System*, *5*, 1–15.

Raiffa, H. (1961). Risk, Ambiguity, and the Savage Axioms: Comment. *The quarterly journal of economics*, *75*, 690–694.

Rizzo, P., Veloso, M. M., Miceli, M., & Cesta, A. (1999). *Goal-based personalities and social behaviors in believable agents*, volume 13.

Roberts, M., Shivashankar, V., Alford, R., Leece, M., Gupta, S., & Aha, D. W. (2016). Goal Reasoning, Planning, and Acting with ActorSim, The Actor Simulator. *Advances in Cognitive Systems*, (pp. 1–16).

Russell, S. J. (1997). Rationality and intelligence. *Artificial Intelligence*, *94*, 57–77.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley.

Shivashankar, V., & Alford, R. (2013). Hierarchical Goal Networks and Goal-Driven Autonomy: Going where AI Planning Meets Goal Reasoning. *Proceedings of the 2013 Annual Conference on Advances in Cognitive Systems: Workshop on Goal Reasoning* (pp. 95–110). Baltimore, MD.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The quarterly journal of economics*, *69*, 99–118.

Simon, H. A. (1996). *The Sciences Of The Artificial*. Cambridge, MA: The MIT Press, third edition.

Steunebrink, B. R., Dastani, M., & Meyer, J.-J. C. (2007). *Emotions as Heuristics for Rational Agents*. Technical report, Department of Information and Computing Sciences, Utrecht University, Utrecht.

Sumerlin, J. R. (1995). Adaptation to Homelessness: Self-actualization, Loneliness, and Depression in Street Homeless Men. *Psychological reports*, *77*, 295–314.

Von Neumann, J., Morgenstern, O., & Others (1944). Theory of games and economic behavior.

Wold, H., Shackle, L. S., & Savage, L. J. (1952). Ordinal Preferences or Cardinal Utility? *Econometrica*, *20*, 661–664.

Zafirovski, M. (2005). Is Sociology the Science of the Irrational? Conceptions of Rationality in Sociological Theory. *The American Sociologist*, *36*, 85–110.