ISSN 0147-6882, Scientific and Technical Information Processing, 2019, Vol. 46, No. 6, pp. 404–415. © Allerton Press, Inc., 2019.
Russian Text © The Author(s), 2018, published in Iskusstvennyi Intellekt i Prinyatie Reshenii, 2018, No. 2, pp. 21–35.

Goal Setting and Behavior Planning for Cognitive Agents

A. I. Panov[a, b, *]

[a]*Artificial Intelligence Research Institute, Federal Research Center "Computer Science and Control," Russian Academy of Sciences, Moscow, 119333 Russia*
[b]*Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Oblast, 141701 Russia*
**e-mail: pan@isa.ru*

Abstract—This paper considers a sign-based approach to the problem of modeling the goal-setting process and integrating it with behavior planning methods. The GoalMAP algorithm, which is based on a psychologically plausible sign-based world model of a cognitive agent, is proposed. The algorithm implements an iterative hierarchical planning process with emphasis on the stage of setting or selecting a new goal. The complexity of the behavior planning algorithm is estimated. Model experiments with a software implementation of the constructed algorithms that demonstrates key features of the adopted approach are performed.

Keywords: cognitive agent, planning, sign, sign based world model, activity theory, cultural-historical approach, goal-setting, behavior planning, GoalMAP

DOI: 10.3103/S0147688219060066

INTRODUCTION

The issue of modeling the goal-setting process in intelligent systems is critical for the task of increasing the degree of autonomy of robotic systems. This issue is addressed both within the classic artificial intelligence research area [1–3] and within new directions of modeling human cognitive functions [4, 5]. Expanding the areas of applicability of robotic systems leads to an expanding range of possible situations that a robot may have to respond to. It is becoming increasingly difficult to foresee and embed information about the possible scenarios involved in achieving a goal that has been set for a robot. There are many examples of situations where a robot or an agent cannot achieve the initial goal and needs to develop a new target or select one from an existing list. One such example is an autonomous deep submergence vehicle that has the goal of exploring a particular area of the seabed. The vehicle may find itself in a situation where an extraneous object (a shipwreck) is found in this area, which renders the initial mission impossible. Goal-setting methods should allow the agent to terminate the previous plan and set a new goal, for example, sending a message about the presence of unidentified objects on the seabed to the command center.

The importance of solving the goal-setting process modeling problem has been emphasized for a long time not only by artificial intelligence specialists, but also by psychologists, in whose area the phenomenon of goal-setting (or goal formulation) is still understudied: "To mathematical models enthusiasts we suggest using the goal formulation phenomenon as a criterion for assessing the quality of proposed mathematical models of mental activity: if the model recreates the process of goal formulation, it is "good," if not, then it is "bad"" [6].

Since examples of the goal-setting process that is available for detailed study are observed only in humans, the achievements of the psychological science in this area should serve as the primary source of theories and experimental data for building goal-setting models. In this article we rely on Russian research in the field of activity theory and the cultural-historical approach [7, 8], which postulate that goal-setting is an integral part of human activity or a separate kind of it, so modeling it should be connected in some way to modeling behavior planning.

A goal as it is understood in artificial intelligence (some final situation described formally) is only one instance of the complex psychological process of operations with goals. Within the framework of the activity theory, several mechanisms of goal-setting are distinguished [6]. Among them are: transformation of the side result of an action into a goal by becoming aware of it and associating it with a motive; reformulation of goals in cases when the originally anticipated result is not achieved; choosing one goal out of a set of goals; formation of a hierarchy and a time sequence of goals; and others. This paper considers a model of one of these goal-setting mechanisms, which is associated with the planning stage, in particular with identifying intermediate goals as an obstacle function and internalizing the given goal by linking it to the motive. This paper presents a psychologically plausible behavior planning algorithm for a cognitive agent and proposes a modification for it aimed at implementing the goal setting mechanisms described above.

In modern research on the topic of goal-setting and, in more general terms, goal reasoning, two main directions can be distinguished: the cognitive direction and the traditional direction. Both approaches define a goal as a class of external environment states that is described by the agent based on some way of knowledge representation and satisfies certain conditions. Within the cognitive direction all operations with goals fall under so-called metacognitive processes that regulate the functioning of other cognitive processes, such as planning.

Within the traditional approach to the topic of goal reasoning, formal definitions of a goal are introduced from the standpoint of planning theory and several types of reasoning that use the concept of a goal are identified. In this regard, research on the so-called goal-driven autonomy [2, 9, 10] should be primarily noted. It highlights four main steps of operations with goals: monitoring for mismatches between knowledge and observations, explaining the mismatch, goal generation, and goal management. The first step involves observing a so-called anomaly or a special event, which may be an observation of the impossibility of achieving the current goal, new contingencies that affect the implementation of the plan, etc. During the second step, the agent can use inductive inference to replenish information about the observed event, for example, by using abductive reasoning or accumulated experience in the form of an event-goal representation. In this case, the goals are assumed to appear and change in the process of the execution of the plan. That is an important point, because the goal-driven autonomy approach does not presuppose the goal-setting stage until the moment when the agent begins to implement the plan and might face a new unforeseen situation. In other words, in this approach some initial goal is in any case set externally, which narrows the very meaning of the goal-setting process.

Let us consider the cognitive direction in goal-setting modeling on the example of cognitive architectures, in which the aim is to build a system of psychologically or biologically plausible control of agent behavior. Some architectures that consider the concept of a goal in one of their subsystems include a structure of a list of goals [11], where the goals are placed according to certain rules and from which they are selected for execution, which can be considered basic operations with goals. In other architectures, part of the episodic memory is used to build predictions of future events and supports the ability to change the goal [12, 13]. Some cognitive architectures [14] attempt to build a separate meta-cognitive level of decision-making, which contains both a mechanism for identifying hypothetical causes of events based on memories of earlier observed states of the environment and an analysis of contingencies based on the identified causal relations. Despite the careful consideration of the links between the subsystem of nomination and prioritization of goals and the other subsystems of the behavior management system (memory, motor-perceptive mechanisms, the subsystem of motivation and evaluation), the approach developed in cognitive architectures remains purely theoretical. Only simple lists of goals and implementations of individual cases, similar to the described goal-driven autonomy, are practically implemented.

This paper is organized as follows. Section 1 provides a formal description of a sign-based world model, which is the basis for the psychologically plausible method of behavior planning for a cognitive agent proposed in Section 2. At the end of Section 2, the complexity of the plan synthesis algorithm is estimated. Section 3 presents two models that implement two types of goal-setting: empirical ("internal") and scenario type ("external"). Section 4 presents experimental results that demonstrate the characteristic features of the presented algorithms, despite being mainly model-theoretic in nature. Some of the results are also discussed in this section.

1. THE WORLD MODEL OF A COGNITIVE AGENT

This paper uses a sign-based world model [4, 15, 16] and its network formalization as described in [17, 18]. Let us introduce the basic concepts and definitions. The main element of the world model is a sign, which can represent both a static object and an action. The sign is defined by its name and contains the following components: image, significance, and personal meaning. The image component encodes the characteristic features of the represented object or process. The significance component represents the scenarios of object use that are available to the team of agents. The personal meaning component defines the role the object plays in the actions a subject can perform on this object. Personal meanings of a sign are formed in the process of the subject's activity and constitute specifications of the scenarios from this sign's significances. Personal meanings expose the activity subject's preferences, express the motive and the sentiment of actions, form the experience of action.

Let us introduce a special structure, that is, a causal matrix, which will be used for constructing formal descriptions of a sign's components.

Definition 1. *The term causal matrix will be used to mean the following structure* $z = \langle e_1, e_2, \ldots, e_t \rangle$ *is a tuple of length* t *containing events* e_i *are bit vectors (columns) of length* h*. Each index* j *of an event vector* e_i *(matrix row* z*) will be assigned a possibly empty tuple of causal matrices* Z_j *such that* $z \notin Z_j$*. Special procedure* $\Lambda : 2^Z \to 2^{\mathbb{N}} \times 2^{\mathbb{N}}$ *assigns two disjoint subsets of column indices* $I^c \subset \mathbb{N}$, $\forall i \in I^c\ i \leq h$ *and* $I^e \subset \mathbb{N}$,

$\forall i \in I^e$ $i \le h$ such that $I^c \cap I^e = \varnothing$ to a subset of causal matrices $Z' \subseteq Z$ from the entire set Z. The set I^c for matrix $z$will be called indices of condition columns, and the set I^e will be called indices of effect columns of matrix z.

The causal matrix acts as the main constituent of the sign components. The structure defined above makes it possible to encode both static information and features of the object and dynamic processes the same way. The built-in option to specify causes and effects makes it possible to encode basic relations identified on the data about the external environment, that is, causal relations.

Definition 2. *The term sign will be used to mean the quadruple $s = \langle n, p, m, a \rangle$, where n is the name of the sign and $p = Z^p$, $m = Z^m$, $a = Z^a$ are tuples of causal matrices called, respectively, image, significance, and personal meaning of sign s .s identified on the data about the external environment—causal relations.*

Based on definitions 1 and 2, the entire set of causal matrices can be split into three subsets (of images, significances, and personal meanings of signs), which are organized into so-called causal networks.

Definition 3. *A causal network $W = \langle V, E \rangle$ is a marked directed graph, where:*

- *Each node $v \in V$ is assigned a tuple of causal matrices $Z^p(s)$ of the image of a sign s, which will be designated as $v \to Z^p(s)$, directed edge $e = (v_1, v_2)$ belongs to the set of directed edges of graph E, if $v_1 \to Z^p(s_1)$, $v_2 \to Z^p(s_2)$ and $v_1 \to Z^p(s_1)$, $v_2 \to Z^p(s_2)$, i.e., if sign s_1 is an element of image s_2;*
- *each directed edge of the graph $e = (v_1, v_2)$, $v_1 \to Z^p(s_1)$, $v_2 \to Z^p(s_2)$ is assigned a mark $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3)$, which is a tuple of three positive integers:*

—ϵ_1 *is the index of the source matrix in tuple $Z^p(s_1)$, it can take on a special significance 0 if any matrix from the tuple can act as a source,*

—ϵ_2*is the index of the target matrix in tuple $Z^p(s_1)$, a row from which is assigned to feature s_1;*

—ϵ_3 *is the index of the column from the target matrix, where the row assigned to feature s_1 has the significance 1, it can take on positive significances (condition columns) and negative significances (effect columns).*

The causal network serves as the basis for identifying basic relations on the set of a sign's components [17] and some processes that model basic cognitive functions (generalizations, significance closures, and agglutinations).

Definition 4. *The term sign-based world model will be used to mean a semiotic network $\Omega = \langle W_m, W_a, W_p, R, \Theta \rangle$, where W_m, W_a, W_p are causal networks on significances, personal meanings, and images respectively, $R = \langle R^m, R^a, R^p \rangle$ is a collection of relations on components of the sign, Θ is a collection of operations on the set of signs.*

The planning process within a sign-based world model can be described through the concepts of activity in a semiotic network and the process of its distribution. Let us introduce an activity label for the causal matrices of network W_x $(x \in \{p, m, a\})$. A set of matrices Z_x^* that possess this label will be called active. The process of propagation of activity constitutes a change in the composition of the set Z_x^* over time (each discrete moment) and is described for each type of causal network by its function: $\varphi_a, \varphi_m, \varphi_p$. The activity propagation process is iterative, i.e., at each step a new composition of the set of active matrices is generated on the basis of the previous composition by adding new matrices connected by directed edges of the network with the current active ones. As the simplest case, we will consider a process in which none of the matrices affect the activity propagation from any other matrix, therefore we can assume that functions φ_x take one active matrix as input and return a new subset of active matrices.

Due to the fact that the edges of the causal networks have directions, we will distinguish between activity propagation in the upward direction through the network, when outgoing directed edges are used (function $\varphi_x^{\uparrow}$), and activity propagation in the downward direction through the network, when ingoing directed edges are used (function $\varphi_x^{\downarrow}$). In the future, for describing the planning algorithm, only functions on the significances and personal meanings networks will be needed. Each function $\varphi_x^{\uparrow}, \varphi_x^{\downarrow}$ will be parameterized by depth of activity propagation d_x, which indicates the depth of the observed directed edges in this direction (upwards or downwards).

2. BEHAVIOR PLANNING OF THE COGNITIVE AGENT

The behavior planning algorithm of an agent is based on the following fundamental principles of the activity theory. First, the proposed algorithm implements hierarchical planning, which separates actions and operations and has a hierarchy of goals. Each action has its own operational composition, which is either known in advance (as components of its subaction image) or obtained as a result of a separate planning algorithm (one of the options of goal-setting, see Section 3). Secondly, actions are always objective, i.e., associated on the significances network with signs that represent objects that perform a particular role in the action. Dividing the set of procedures available to the agent into actions aimed at achieving the goal and operations that reflect the specifics of a particular situation is also consistent with Kahneman's theory of fast and slow thinking [19].

The basic idea of planning in a sign-based world model comes down to using the locality of the organization of knowledge, which is achieved by the fact that procedural and declarative knowledge is grouped into signs on the basis of experience of interacting with some object or performing actions in some situation. Since the agent's goal-setting activity is objective and situational, i.e., conditions of actions are determined by the current situation, then combining information about the key features of the object, the actions in which it can take part, and the experience of interacting with it that the agent has already accumulated makes it possible to localize or limit the search for the information necessary for determining the next action.

Let us first describe the usual planning algorithm in a sign-based world model [18]; in the next section we will focus on the goal-setting stage, one of the stages of this algorithm. We will consider the case of symbolic planning, which doesn't involve the tasks of recognizing objects and situations of the external environment; this means that the image components of all signs will be empty.

The planning process in a sign-based world model is realized by means of the MAP-algorithm and is performed in a backwards way: from the final situation to the initial one. Let us briefly describe its main stages. The algorithm receives a description of the problem as input:

$$T = \langle N_T, S, Sit_{start}, Sit_{goal} \rangle,$$

Where N_T is the identifier of the task, S is a set of signs of the semiotic network, $Sit_{start} = \langle \varnothing, \varnothing, a_{start} \rangle$ is the sign of the initial planning situation with the meaning $a_{start} = \{z^a_{start}\}$ and empty significance and image, $Sit_{goal} = \langle \varnothing, \varnothing, a_{goal} \rangle$ is the target situation with the meaning $a_{goal} = \{z^a_{goal}\}$ and empty significance and image.

The significances of the signs of the situation are empty because in the case that we are considering no generalized actions from the group to which the agent belongs are associated with the task assigned to the agent, i.e., no solution schemes are developed for this task in the group of agents. In the general case, the task T is the result of the "signification" procedure, that is, forming a world model according to the original descriptions of the planning domain D, which specifies lists of possible actions and object types, and the planning problem P, which includes the definitions of the starting conditions and the final goal (step 1 of algorithm 1):

```
Input: description of the planning domain $D$, description of the planning problem $P$, maximum depth of iteration $i_{\max}$
Output: plan $Plan$
1: $T = \langle N_T, S, Sit_{start}, Sit_{goal} \rangle := \mathrm{GROUND}(P)$
   //$N_T$—task identifier, $S$—set of signs, $z_{cur} := (e$ —initial situation with meaning $a_{start}$,
   $Sit_{gool} = \langle id_{gool} \varnothing, \varnothing, \{z^a_{gool}\} \rangle$—target situation with meaning $a_{gool}$
2: $Plan := \mathrm{MAP\_SEARCH}(T)$
3: function $\mathrm{MAP\_SEARCH}(T)$
4: $z_{cur} := z^a_{goal}$
5: $z_{start} := z^a_{start}$
6: $Plans := \mathrm{MAP\_ITERATION}(z_{cur}, z_{start}, \varnothing, 0)$
7: $\{Plan_0, Plan_1, \ldots\} = \mathrm{SORT}(Plans)$
8: return $Plan_0$
```

Algorithm 1. A MAP algorithm of behavior planning: general scheme.

The MAP-algorithm results in $Plan = \{\langle z^a_{s1}, z^a_{p1} \rangle, \langle z^a_{s2}, z^a_{p2} \rangle, \ldots, \langle z^a_{sn}, z^a_{pn} \rangle\}$, that is, a sequence of length n pairs $\langle z^a_{si}, z^a_{pi} \rangle$, where z^a_{si} is the causal matrix of some node of the network on personal meanings that represents the ith planning situation and z^a_{pi} is the causal matrix of some personal meaning that represents the action applied to situation z^a_{si}. Meanwhile, situation z^a_{si+1} is a result of performing action z^a_{pi} in a sense that will be explained later, while considering the algorithm, $z^a_{s1} := z^a_{start}$ is the causal matrix that corresponds to the meaning of the initial situation and $z^a_{sn} = z^a_{goal}$ is the causal matrix that corresponds to the meaning of the target situation.

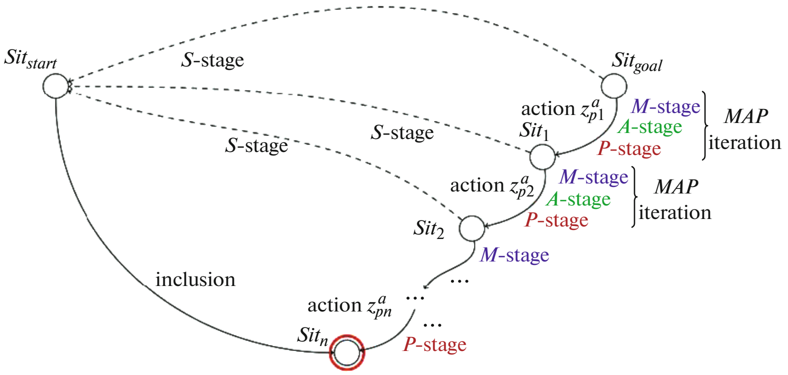

Fig. 1. The general scheme of behavior planning in a sign-based world model. The double circle marks achieving a situation that includes the initial situation (iterations end).

The planning process is hierarchic and consists of repeating the MAP iteration, which consists of four stages (Fig. 1)

- S-stage – the search for precedents of actions performed in the current situation,
- M-stags – the search for applicable actions on the significance network,
- A-stage – the generation of personal meanings that correspond to the found significances,
- P-stage – the construction of a new current situation based on the set of features of the conditions of the found actions.

In short, the MAP algorithm performs an iterative generation of new causal matrices z_{next} of personal meanings on the basis of the current active matrix z_{cur} until the step limit i_{max} is reached (step 10) or the initial matrix z_{start} that corresponds to the personal meaning a_{start} of the initial situation is fully activated (step 41). For the first iteration, the matrix corresponding to the personal meaning of the target situation z^a_{goal} is used as the current active matrix (step 4). After completing all of the iterations, the plans that are found are sorted by length (step 7) and the shortest one is considered to be the solution to the planning problem in the sign-based world model (step 8).

The first stage of the MAP iteration is the S-stage. Its main point is the search for precedents in the world model of the intellectual agent, i.e., the search for actions that have been previously performed in current conditions z_{cur}. This is done by looking through all the signs in the world model S and their personal meanings $a(s)$ (steps 13–16). If current conditions z_{cur} are satisfied by the matrix $z_a \geq z_{cur}$, then the list of precedents $\hat{A}_{case}$ is appended by the results of activity propagation over the personal meanings network from the sign s over distance d_a (step 16). Since directed edges connect causal matrices on the personal meanings network, which represent actions and objects that participate of these actions, the actions that the agent has previously performed with the sign s will be assigned to the matrix activated from the sign s. This is the only global search stage in the MAP algorithm that will be replaced by local search if a full-fledged P-stage is implemented, which would involve a search for situations similar to z_{cur} in some neighborhood of the image of the situation z_{cur}:

The next stage of the MAP algorithm is the M-stage, at which activity propagates over the network of personal meanings over distanced_a in order to activate all the signs associated with the current situation (step 17). The elements of the resulting set of causal matrices A^* serve as starting points for activity propagation across the significances network: for each matrix z_a the necessary node on the causal significances network is determined by the binding function Ψ^m_a, and from that node the activity propagates over distance d_m (step 20). If the activated matrices are causal, they are added to the set of active significances M^* (step 22). As significances define the roles of the actions and concatenate classes and subclasses, performing the procedure $\varphi^{\uparrow}_m$ results in actions performed with the object $s(z_a)$ or its superclasses being added to the set M^*. Both steps of activity propagation on networks ($\varphi^{\downarrow}_a$ and $\varphi^{\uparrow}_m$) define the area of local search in the world model:

```
9: function MAP_ITERATION($z_{cur}, z_{start}, Plan_{cur}, i$)
10: if $i \geq i_{max}$ then
11:     return $\varnothing$
```

```
12: $\hat{A}_{case} := \varnothing$ // List of precedents
    // S-stage
    // Search for precedents of actions performed in current conditions
13: for all $s \in S$ do
14:   for all $z_a \in a(s)$ then
15:     if $z_a \geq z_{cur}$ then
16:       $\hat{A}_{case} = \hat{A}_{case} \bigcup \varphi_a^{\uparrow}(s, d_a)$
```

Algorithm 2. A MAP algorithm of behavior planning: S-stage.

```
    // M-stage
    // Activity propagation downwards on the personal meanings network
17: $A^* = \varphi_a^{\downarrow}(z_{cur}, d_a)$
18: $M^* = \varnothing$
19: for all $z_a \in A^*$ do
      // Activity propagation upwards on the significances network
20:   for all $z_m \in \varphi_m^{\uparrow}(s(z_a), d_m)$ do
21:     if $I^c(z_m) \neq \varnothing$ them
22:       $M^* := M^* \bigcup \{z_m\}$
```

Algorithm 3. MAP algorithm of behavior planning: M-stage.

Next, we proceed to the A-stage, at which causal matrices are generated on the network of personal meanings, whose are actions specified with respect to the current conditions z_{cur} and determined by the active significances from the set M^*. Steps 25–27 serve this purpose. They involve activity propagation on the causal significances network over distance d_m, which leads to the formation of a set $\hat{M}^*$ of signs associated with the role structure of the procedural matrix z_m, i.e., the set includes objects that can potentially replace roles in z_m actions. Next, using the binding function Ψ_m^a, a new causal matrix is generated on the personal meanings network. It copies the z_m^* significance with replacement of abstract signs-roles with object signs linked to the roles by class-subclass relations. The A-stage then involves selecting such causal matrices that represent actions that are feasible under the current conditions z_{cur} (steps 29–32). To do this, all causal matrices, the effects z_{shift} of which are not included in the current situation $z_a \not\geq z_{shift}$, are removed (we recall that planning is performed in the backwards direction).

The A-stage concludes with performing one of the operations in the world model θ_a, which, in this case, performs the meta-regulation function, checking on some heuristics that express, for example, the rule that actions cannot be repeated, or that actions that brings the situation closer to the initial conditions z_{start} the fastest are preferred (step 33). This allows further reduction in the search space. Any heuristic rule can also be represented as a causal matrix of the personal meaning of the sign, which represents the internal strategy of its behavior planning:

```
    // A-stage
23: $\hat{A}_{gen} = \varnothing$
24: for all $z_m \in M^*$ do
      // Activity propagation downwards on the significances network
25:   $\hat{M}^* = \varphi_m^{\downarrow}(z_m, d_m)$
26:   for all $z_m^* \in \hat{M}^*$ do
27:     $\hat{A}_{gen} := \hat{A}_{gen} \bigcup \{\Psi_m^a(z_m^*)\}$
      // Combining the activity of the created meanings and the current situation
```

```
28: $\hat{A} = \hat{A}_{gen} \cup \hat{A}_{case}$
29: **for all** $z_a \in \hat{A}$ **do**
30: $z_{shift} = (e_i | i \in I^c)$
31: **if** $z_{cur} \ngeq z_{shift}$ **then**
32: $\hat{A} = \hat{A} \backslash \{z_a\}$
    // Metacognitive heuristics check
33: $\hat{A} = \{\theta_a(z_a) | z_a \in \hat{A}\}$
34: **if** $\hat{A} = \varnothing$ **then**
35: **return** $\varnothing$
```

Algorithm 4. MAP algorithm of behavior planning: A-stage.

The MAP algorithm concludes with the P-stage. At this stage, for each generated causal matrix z_a that represents some action, a new situation Sit_{next}, which is the result of a reverse performance of the action in the current conditions z_{cur}, is formed. The reverse performance (step 39) involves forming a causal matrix z_{next} that consists of events that either are columns-conditions of the action $e_i \in \{e_k | e_k \in z_a, \ k \in I^c(z_a)$ or belong to the current active causal matrix and are not columns-effects of the action $e_i \in z_{cur} \wedge e_i \notin \{e_j | e_j \in z_a, \ j \in I^e(z_a)\}$. Since the case considered here is one of symbol planning, the P-phase is shortened and does not include the process of constructing an image of the new situation through the use of the binding function Ψ_m^a. Image construction would allow narrowing the precedents search at the S-stage:

```
    // P-stage
36: $Plans_{fin} := \varnothing$
37: **for all** $z_a \in \hat{A}$ **do**
38: $Plan_{cur} = Plan_{cur} \cup \{\langle z_{cur}, z_a \rangle\}$
    // Generating a new situation; executing the action
39: $z_{next} := (e_i | (e_i \in z_{cur} \wedge e_i \notin \{e_j | e_j \in z_a, \ j \in I^e(z_a)\}) \vee e_i \in \{e_k | e_k \in z_a, \ k \in I^c(z_a)\})$
40: $Sit_{next} = \langle id_{next}, \varnothing, \varnothing, \{z_{next}\} \rangle$
41: **if** $z_{next} \geq z_{start}$ **then**
42: $Plans_{fin} = Plans_{fin} \cup \{Plan_{cur}\}$
43: **else**
44: $Plans_{rec} :=$ MAP_ITERATION$(z_{next}, z_{start}, Plan_{cur}, i + 1)$
45: $PlansLfin = Plans_{fin} \cup Plan_{rec}$
46: **return** $Plans_{fin}$
```

Algorithm 5. MAP algorithm of behavior planning: P-stage.

Applicable action $\langle z_{cur}, z_a \rangle$ is added to the current $Plan_{cur}$. If the new situation does not cover the initial situation (step 43), the iterations continue with the new current situation, appending the set of all generated plans $Plans_{fin}$.

The constants d_a, d_m that determine the depth of activity propagation in causal networks are algorithm parameters and define the internal characteristics of the world model host and differing from agent to agent. In model experiments, these parameters usually do not exceed five. Increasing the significances of these constants leads to a significant increase in the complexity of the MAP algorithm, in accordance with Theorem 1.

Theorem 1. The complexity of a MAP_ITERATION iterative step of the MAP algorithm in a symbol setup (algorithms 1–5) in the worst case scenario equals $O(N + C_1^{d_a} C_2^{2d_m})$, where N is the number of signs in the world model; C_1, C_2 are some constants; d_a, d_m are algorithm parameters that determine the depth of activity propagation in the causal networks of personal meanings and significances respectively.

Proof. Let the complexity of the procedure of activity propagation over causal networks be limited by the constant c_{φ}, the complexity of the check for nestedness of causal matrices be limited by the constant c_{in}, the complexity of the generation of a new matrix be limited by c_{next}, the complexity of the operation of the metacognitive procedureθ_a be limited by c_{meta}, and the complexity of the operation of binding functions Ψ be limited by c_{Ψ}. Then, the complexity of the S-stage (steps 13–16) is limited by the significance $c_{\varphi}c_{in}N$, the complexity of the M-stage (steps 17–22) – by the significance $c_{\varphi}N_a^{\downarrow d_a}N_m^{\uparrow d_m}$, where $N_a^{\downarrow}$ is the maximum number of outgoing edges for a node on the personal meanings network and $N_m^{\uparrow}$ is the maximum number of incoming edges of a node on the significances network. Since the potency of the set M^* is at most $N_a^{\downarrow d_a}N_m^{\uparrow d_m}$, the complexity of the A-stage (steps 23–32) is estimated at $c_{\varphi}c_{\Psi}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$, where $N_m^{\downarrow}$ is the maximum number of outgoing edges for a node on the significances network. The metacognitive check (step 33) adds the significance $c_{meta}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$ to the complexity, because the potency of the set $\hat{A}$ is at most $N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$. Finally, the complexity of the P-stage is limited by the significance $c_{next}c_{\varphi}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$ and the significance $N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$ specifies the degree of branching of the entire recursive MAP algorithm. Summing up, the complexity of the entire MAP_ITERATION iterative step is:

$$\begin{gathered}
O(c_{\varphi}c_{in}N + c_{\varphi}N_a^{\downarrow d_a}N_m^{\uparrow d_m} + c_{\varphi}c_{\Psi}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m} \\
+ c_{meta}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m} + c_{next}c_{\varphi}N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}) \\
= O(c_{\varphi}c_{in}N + (c_{\varphi}c_{\Psi} + c_{meta} \\
+ c_{next}c_{\varphi})N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}) < O(N + C_1^{d_a}C_2^{2d_m}),
\end{gathered}$$

where $C_1 = N_a^{\downarrow}$, $C_1 = \max(N_m^{\downarrow}, N_m^{\uparrow})$.

The complexity of the entire MAP algorithm depends on the degree of branching of the MAP_ITERATION procedure, which can be estimated from above as $N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m}$ (the power of set $\hat{A}$) and in general on the depth of the resulting recursion, which, in turn, is determined by the task. The overall complexity of the algorithm is thus estimated as $O((N_a^{\downarrow d_a}N_m^{\uparrow d_m}N_m^{\downarrow d_m})^r(N + C_1^{d_a}C_2^{2d_m}))$, where r is the maximum depth of recursion.

It should be noted that if a sub-stage for the construction of a current situation image is considered at the P-stage, instead of the full number of signs N, the search at the S-stage would be carried out only on a certain subset with the potency on the order of $N_p^{d_p}$, where N_p is the maximum number of edges of a node on the images network, and d_p is the depth of activity propagation on the images network.

3. THE GOAL-SETTING STAGE

As mentioned in the previous paragraph, the result of constructing a plan is a chain of actions, a sequence of causal matrices that makes it possible to get from the initial state of the environment to the target state. The found and possibly executed chain of actions can be stored in the agent's world model as a sign $s_{plan} = \{p_{plan}, \varnothing, a_{plan}\}$ with the following components: the image of the plan $p_{plan} = \{z^p_{plan}\}$ that includes a causal matrix z^p_{plan} that consists of columns, each containing only one link to the sign of the action that is performed at the current step of the plan. Thus, if $Plan = \{\langle z^a_{s1}, z^a_{p1}\rangle, \langle z^a_{s2}, z^a_{p2}\rangle, \ldots, \langle z^a_{sn}, z^a_{pn}\rangle\}$, then $z^p_{plan} = (e_1, e_2, \ldots, e_n)$, where n is the number of steps of the plan and the column $e_i = \langle 0, \ldots, 0, s_{pi}, 0, \ldots, 0\rangle$ contains only one link to the sign of the action s_{pi} from the i-th step of the plan. The causal matrix of the personal meaning of the plan $a_{plan} = \{z^a_{plan}\}$ contains two columns with links to the signs of the initial situation $s_{s1} = Sit_{start}$ and the target situation $s_{sn} = Sit_{goal}$. Thus, the agent in the general case and as operating experience retains information about the conditions and effects of actions and about their operational composition.

In some situations, it is possible to retain only the meaning of the plan without its operational composition:

- in coalition cases, when the plan is created by a team of agents and different agents, not just the host of the world model, can be subjects of actions [20];
- in cases where the agent's resources are limited, which is especially relevant to compact hardware robotics systems;
- in cases when the execution of the plan has negative results from the tactical level of agent control, for example, as a result of peculiarities of constructing movement trajectories for a group of agents [21].

Such partial steps with the missing image and operating component will be called *schematic*. The agent's operating experience is used at the S-step of the planning algorithm and will also be used in the goal-setting process.

We distinguish two types of goal-setting: the empirical type and the scenario type (Fig. 2). The first type is characterized by the use of experience in performing actions in similar planning situations, and the second is characterized by the use of known scenarios for achieving the goal, which are stored on the significances network of the agent's world model.

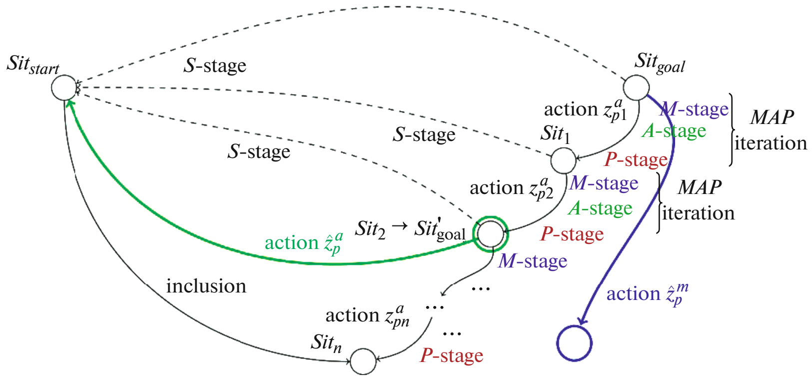

Fig. 2. The goal-setting stage of the MAP planning algorithm. The heavy lines mark the new steps of selection "schematic" actions $\hat{z}^a_p$ and $\hat{z}^m_p$

Empirical goal-setting is an addition to the P-stage of the MAP planning algorithm and allows one to shorten the process of constructing the current plan, set a new goal (in this case, a sub-goal), and move on to constructing a new plan of achieving that new goal.

Additional steps in the GoalMAP algorithm (see Algorithm 6) include checking that the found action is schematic (i.e., its figurative component, operational composition are empty) and running the new *MAP_SEARCH* planning algorithm. Once it has been established that it is possible to perform the currently considered action from the initial situation $z_{next} \geq z_{start}$, $Plan_{cur}$ is considered to be constructed and is added to the set of current plans $Plan_{fin}$. However, the last added action z_a (the corresponding transition with action $\hat{z}^a_p$ is marked in Fig. 2 by a heavy line) might be schematic, i.e., the components of its action are not known and its image $\Psi^p_a(z_a)$ is empty. In this case, the previous planning situation z_{cur} becomes the new goal and constitutes the personal meaning of the new target situation $Sit_{new} = \langle \varnothing, \varnothing, z_{start} \rangle$. In other words, the constructed plan, despite being complete (all the steps of the transition from the initial situation to the target situation are known), is not executable, since the operational composition is not known for every action. It is to clarify this issue that a new goal is formed and the planning process is started again, with no changes to the initial situation:

```
if $z_{next} \geq z_{start}$ then
    $Plans_{fin} = Plans_{fin} \cup \{Plan_{cus}\}$
    * if $\Psi^p_a(z_a) = \varnothing$ then
    *     $MAP\_SEARCH(\langle \text{“}sub\text{”} + N_T, S, Sit_{start}, \langle \varnothing, \varnothing, \{z_{cur}\} \rangle \rangle)$
else
    $Plans_{rec} := \mathrm{MAP\_ITERATION}(z_{next}, z_{start}, Plan_{cur}, i + 1)$
    $Plans_{fin} = Plans_{fin} \cup Plans_{rec}$
```

Algorithm 6. GoalMAP algorithm of behavior planning: P-stage
Asterisks indicate additional to the original MAP algorithm P-stage steps.

Let us consider an example: the process of planning actions for achieving the goal "being in St. Petersburg" from the initial situation "being in Moscow." At some stage, the following chain of actions can be found: "buy a train ticket," "arrive at the Leningradsky railway terminal," "travel by train," "arrive at the Moskovsky railway station." In this chain, the final (from the reverse planning point of view) action is executable in the initial situation, but is schematic, i.e., it is known that the subject has previously bought train tickets, but the specific operational composition of this action is either unknown or not considered in the current plan and can be refined in the course of a separate planning process with the new target situation "having a train ticket."

In contrast to the empirical approach, scenario type goal-setting occurs before the main MAP planning algorithm is launched (the corresponding transition is marked in Fig. 2 with the right heavy line). In general, the MAP algorithm is used for initial and target situations in which the conditions are set by specific objects (certain “a” and “b” blocks) and relations on their set that are specific to only this case (block “a” specifically lies on top of block “b”). However, in some cases planning can be performed in an abstract domain, for example when the conditions of situations are defined not by objects, but by their classes (how to move a set of any blocks from one position to another), and the situation is standard for the group to which the agent belongs. In this case, it is not required to generate personal meanings in order to define an applicable action and a new current situation on the significances network can be built using a scheme of the action instead of the action itself. The resulting situation becomes the new goal and the plan to achieve it is based on a separate planning process. The current plan ends up consisting of one generalized action and may need to be adjusted and specified when it comes to executing it:

```
function MAP_SEARCH($T$)
$z_{cur} := z^a_{goal}$
for all $z_m \in \varphi^{\uparrow}_m(Sit_{goal}, d_m)$ do
    ** $z_a = \Psi^a_m(z_m)$
    ** if $I^e(z_a) \neq \varnothing \wedge \exists k : k \in I^e(z_a),\ e_k \in z_a,\ Sit_{goal} \in e_k$ then
        ** $z_{cur} := (e_i | e_i \in z_a \wedge e_i \notin \{e_j | e_j \in z_a, j \in I^e(z_a)\})$
$Plans := \text{MAP\_ITERATION}(z_{cur}, z_{start}, \varnothing, 0)$
$\{Plan_0, Plan_1, \ldots\} = \text{SORT}(Plans)$
```

Algorithm 7. GoalMAP algorithm of behavior planning: a scenario-type case. Double asterisks indicate additional to the original MAP algorithm steps.

Additional steps in the GoalMAP algorithm (see Algorithm 7) include looking through possible procedural matrices z_m that belong to the significances of a sign of the target situation Sit_{goal} or one of its classes (the result of activity propagation $\varphi^{\uparrow}_m$). The corresponding significances of causal matrices of personal meanings (specified actions in the current context) are checked for possible solutions to the current problem. Since the target situation may be standard, there may be actions in which it is involved as an effect $Sit_{goal} \in e_k, k \in I^e(z_a)$. Conditions of the found actions constitute the new target z_{cur}, for which the current plan *MAP_ITERATION* will be constructed. It should be noted that a situation is typical for a group of agents not only in cases when it involves generalized concepts, but also when a frequently encountered situation is coordinated within the group of agents and becomes typical for that group.

One example of this type of goal-setting is reasoning on achieving the goal “becoming famous.” Even subjects who have no experience in performing specific actions in this direction are familiar with the scenario adopted in their cultural group that says that if one wants to be famous it is common practice to become a writer, i.e., to write novels. Thus, the subject uses the initial situation “being a writer” of the action “becoming famous” as a new target situation for a new process of behavior planning.

4. MODEL EXPERIMENTS

Let us consider model-theoretic examples in order to showcase both types of goal-setting (the empirical and scenario types). The first case will demonstrate how the experience of solving a problem is retained in a schematic form to later serve as the basis for launching the planning procedure for a new goal. In the second case, a new scenario of achieving the goal that is applicable to the new task and usable for generating a new goal is added to the agent’s world model in advance. Both model experiments will be carried out in the widespread “blocks world” planning domain, which has been used to demonstrate the operation of the main MAP algorithm as well [18]. The description of the domain in PDDL [22] consists of the definition of the type “block,” four predicates (“ontable,” “clear,” “holding,” “armempty”), and four actions (“unstack,” “stack,” “pickup,” “putdown”). The domain description is translated into the agent’s sign-based world model, in which a sign is created for each listed component and corresponding causal matrices are constructed on the significances network.

In both cases, we consider the task of constructing a tower made of five blocks A, B, C, D, E, which are all initially on the table. Corresponding causal matrices are created for the initial and target situations on the personal meanings network. Figure 3 depicts a fragment of the causal network for the target situation (ABCDE tower).

Let us assume that the agent has previously gained experience in planning to achieve the target situation “tower of four blocks ABCD” from the initial situation

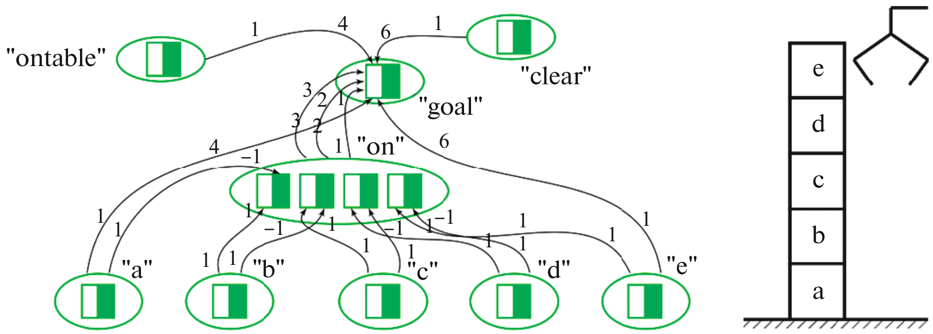

Fig. 3. The target situation of the "tower made of the blocks ABCDE" planning problem (on the left is a fragment of the causal network on personal meanings, on the right is a scheme of the situation). The sign "ontable" and "clear" stand for the corresponding predicates, the sign "goal" stands for the target situation. The numbers on the arrows indicate indices ϵ_1 and ϵ_3 and labels of the directed edges of the causal network (see Definition 3)

"all blocks on the table." Usually, the planning experience is retained in the agent's world model in the form of fragments of the causal network on images and personal meanings. The network on images stores the operational composition of the action (which smaller subtasks does the action "build a tower" consist of) (Fig. 4). The network on personal senses stores information about the initial and final situations for which the plan was created.

Let us go back to building a plan to achieve the new target situation "ABCDE tower." Using the previous planning experience, the agent will create the following plan: "stack block e on block d," "pick block e up from the table," "build the ACBD tower." The last action is represented by a new sign preserved in the world model after the construction (and possible successful implementation) of the plan to achieve the situation "tower ABCD." We suppose that, for example due to lack of memory, the operational structure of this action has not been preserved, i.e., the binding procedure $\Psi_a^p(z_a)$ in Algorithm 6 produces an empty set. In accordance with the algorithm of empirical goal-setting, on completing the formation of the current plan the agent begins planning to reach the new goal "building the ABCD tower." If the agent remembered all its accumulated experience, it is obvious that it would not need to set a new goal in this case.

Now we suppose that the action of building the "ABCDE tower" was either repeated many times and coordinated within some group of agents that the host of the world image belongs to, or information about this action was received by the agent in the form of some message. In this case, the agent only has a scheme of such action in the form of a fragment of a causal network on significances that does not solve the entire problem, but starts with the situation "tower ABCD" (Fig. 5). Having such a scenario lets the agent know that group experience guarantees that the action building the "ABCDE tower" is executable from a situation where there is an "ABCD tower." Therefore, the initial goal is considered to be achieved, and the situation "ABCD tower" becomes the new goal, which is what implements the scenario type of goal setting.

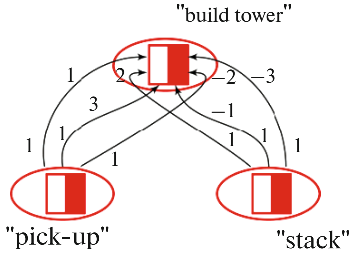

Fig. 4. A fragment of the causal network on images that demonstrates retaining of the experience of building a tower (sign "build tower"). The signs "pick-up" and "stack" stand for the corresponding actions.

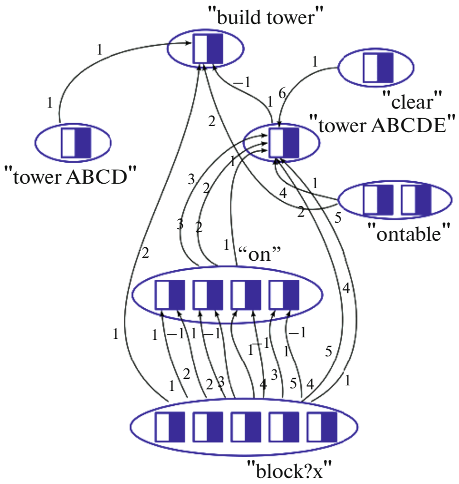

Fig. 5. A fragment of the causal network on significances that a scenario of achieving the goal "ABCDE" tower" from the initial situation "ABCD tower and a cube on the table."

CONCLUSIONS

This article presented an original method of goal-setting that is based on the ideas of operations with goals in the psychological activity theory. It was shown that some goal-setting mechanisms are integrated into the process of behavior planning. GoalMAP, a new behavior planning algorithm that has a goal-setting stage and is based on a sign-based world model was presented. The proposed algorithm was implemented in the form of a software system, with which a number of model experiments that demonstrate the main features of the proposed method were conducted.

FUNDING

This study was supported in part by the Russian Foundation for Basic Research, projects no. 17-07-00281, 16-37-60055.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Alford, R., et al., Hierarchical planning: Relating task and goal decomposition with task sharing, *Int. J. Conf. Artif. Intell.,* 2016, pp. 3022–3028.
2. Molineaux, M., Klenk, M., and Aha, D.W., Goal-driven autonomy in a navy strategy simulation, *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence,* 2010, pp. 1548–1554.
3. Cox, M.T., Perpetual self-aware cognitive agents, *AI Mag,* 2007, vol. 28, no. 1, pp. 32–45.
4. Osipov, G.S., et al., Signs-based vs. symbolic models, *MICAI 2015: Advances in Artificial Intelligence and Soft Computing,* 2015, pp. 3–11.
5. Osipov, G.S., Panov, A.I., and Chudova, N.V., Behavior control as a function of consciousness. I. World model and goal setting, *J. Comput. Syst. Sci. Int.,* 2014, vol. 53, no. 4, pp. 517–529.
6. *Psikhologicheskie mekhanizmy tseleobrazovaniya* (Psychological Mechanisms of Goal Formulation), Tikhomirov, O.K., Ed., Moscow: Nauka, 1977.
7. Leontyev, A.N., *The Development of Mind,* Kettering: Erythros Press and Media, 2009.
8. Vygotsky, L.S., *Thought and Language,* MIT Press, 1986.
9. Muñoz-Avila, H., et al., Goal-driven autonomy with case-based reasoning, *ICCBR 2010: Case-Based Reasoning. Research and Development,* 2010, pp. 228–241.
10. Roberts, M., et al., Iterative goal refinement for robotics, *Working Notes of the Planning and Robotics Workshop at ICAPS,* 2014.
11. Anderson, J.R., *How Can the Human Mind Occur in the Physical Universe?,* Oxford: University Press, 2007.
12. Kurup, U., et al., Using expectations to drive cognitive behavior, *AAAI Conf. Artif. Intell.,* 2011, pp. 221–227.
13. Laird, J.E., *The Soar Cognitive Architecture,* MIT Press, 2012.
14. Samsonovich, A.V., Goal reasoning as a general form of metacognition in BICA, *Biol. Inspired Cognit. Archit.,* 2014, vol. 9, pp. 105–122.
15. Osipov, G.S., Panov, A.I., and Chudova, N.V., Behavior control as a function of consciousness. II. Synthesis of a behavior plan, *J. Comput. Syst. Sci. Int.,* 2015, vol. 54, no. 6, pp. 882–896.
16. Osipov, G.S., Sign-based representation and word model of actor, 2016 *IEEE 8th International Conference on Intelligent Systems (IS),* 2016, pp. 22–26.
17. Osipov, G.S. and Panov, A.I., Relationships and operations in a sign-based world model of the actor, *Sci. Tech. Inf. Process.,* 2018, vol. 45, no. 5, pp. 317–330.
18. Panov, A.I., Behavior planning of intelligent agent with sign world model, *Biol. Inspired Cognit. Archit.,* 2017, vol. 19, pp. 21–31.
19. Kahneman, D., *Thinking Fast and Slow,* New York: Penguin, 2011.
20. Kiselev, G.A. and Panov, A.I., Synthesis of the behavior plan for group of robots with sign based world model, in *Interactive Collaborative Robotics,* Ronzhin, A., Rigoll, G., and Meshcheryakov, R., Eds., Springer, 2017, pp. 83–94.
21. Panov, A.I. and Yakovlev, K.S., Psychologically inspired planning method for smart relocation task, *Procedia Comput. Sci.,* 2016, vol. 88, pp. 115–124.
22. Gerevini, A.E., et al., Deterministic planning in the fifth international planning competition: PDDL3 and experimental evaluation of the planners, *Artif. Intell.,* 2009, vol. 173, nos. 5–6, pp. 619–668.

Translated by A. Ovchinnikova