

УПРАВЛЕНИЕ ПОВЕДЕНИЕМ КАК ФУНКЦИЯ СОЗНАНИЯ. II. САМОСОЗНАНИЕ И СИНТЕЗ ПЛАНА*

© 2015 г. Г. С. Осипов, А. И. Панов, Н. В. Чудова

Москва, Институт системного анализа РАН

Рассматривается семантический уровень описания функций, которые в психологии принято относить к функциям сознания и самосознания. Исследуется механизм работы компонент знака, введённых в первой части статьи. На основе описания знака на семантическом уровне исследуется сходимость основного итерационного процесса образования знака — связывания образной компоненты знака и его значения. Введение алгоритмов работы компонент знака позволяет построить алгоритм процесса синтеза плана поведения, а также построить новую архитектуру интеллектуальных агентов, обладающих, в частности, способностями к распределению ролей в коалициях.

Введение. Связь с первой статьёй.

Правильное введение и плавный переход к образной компоненте.

Как было подробно изложено в первой части статьи [1], в качестве базовых психологических теорий, в которых даётся не только качественное описание свойств когнитивных функций, но и приводятся структурные описания лежащих в их основе психических образований, в предложенном подходе использованы культурно-исторический подход Выготского—Лурии [2, 3], теория деятельности Леонтьева [4] и модель психики Артемьевой [5]. Согласно приведённым теориям высшие сознательные когнитивные функции осуществляются в рамках так называемой мотивированной предметной деятельности, когда объекты и процессы внешней среды опосредованы для субъекта специальными образованиями, называемыми знаками. Процесс задействования знака в той или иной когнитивной функции имеет три образующих: образ, значение и личностный смысл. Образная составляющая отвечает за функции воспроизведения и отличия опосредуемого предмета или процесса в ходе деятельности. Составляющая значения представляет собой место данного знака в той или иной надпсихологической знаковой системе, которая отражает в функциональном смысле наработанные общей исторической практикой коллектива—владельца данной знаковой системы способы использования опосредуемого предмета и процесса. Наконец, составляющая личностного смысла несёт в себе собственный опыт действия субъекта с денотатом знака, который выражается в том числе и в интегральной оценке роли этого денотата в его текущей деятельности: способствует ли данный процесс или объект удовлетворению текущего мотива.

Трёхкомпонентная структура элемента индивидуального знания, которая как было сказано выше, в психологии называется знаком, подтверждается и теми работами нейрофизиологов, в которых предпринимается попытка построить общую теорию работы мозга человека. Так в теории повторного входа Эделмена [6] и Иваницкого [7] утверждается, что образование осознанного ощущения или фиксация входного потока информации происходит только в том случае, когда активированное сенсорным входом возбуждение через ассоциативные зоны коры от гиппокампа, а затем от гипоталамуса накладывается на сенсорный след в проекционной коре. Такой «круг ощущений» 1, проходящий за характерное время в 150-300

*Работа выполнена при поддержке РНФ (грант № 14-11-00692).

мс, последовательно активирует три компоненты индивидуального знания: образную (проекционная и сенсорная зоны коры), компоненту значения (гиппокамп) и личностного смысла (гипоталамус).

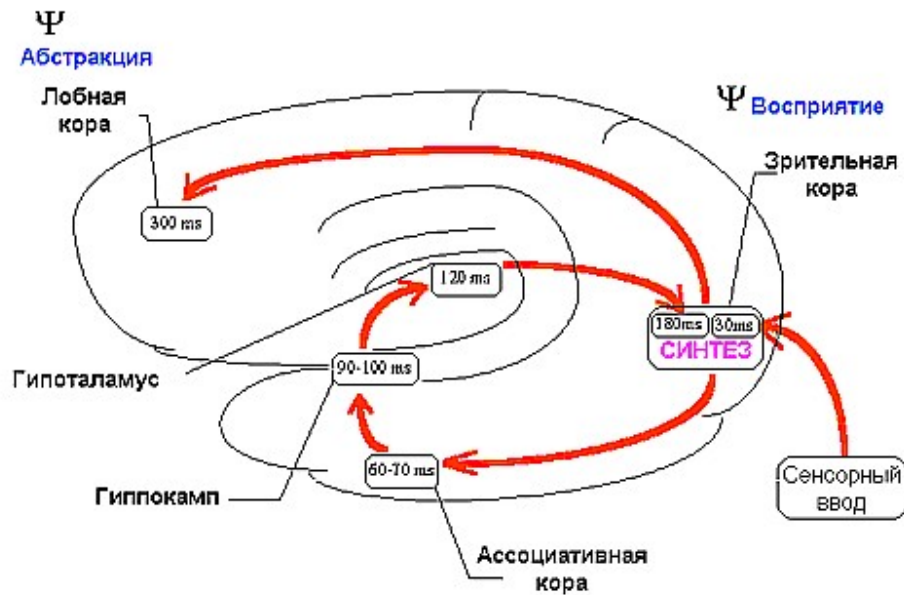


Рис. 1: «Круг ощущений» по Иваниукому (адаптировано из [7]).

Кроме того, по современным нейрофизиологическим представлениям строение коры головного мозга практически однородно во всем своем объеме (наличие колонок некоротекса). При этом связи между достаточно малыми зонами коры (так называемый коннектом), явно указывают на иерархичность ее строения и на присутствие как восходящих, так и обратных, нисходящих связей. Отсюда следует, что компоненты элемента индивидуального знания должны обладать иерархическим однородным строением с восходящими потоками информации и нисходящей обратной связью. Кроме того, образная компонента должна иметь такую функцию распознавания, которая кроме категоризации статических объектов и динамических процессов использует обратную связь для предсказания сигнала в следующий момент времени.

1. Семантический уровень. В качестве математической модели базовой составляющей всех компонент элемента индивидуального знания был предложен следующий бесконечный автомат Мили с переменной структурой и конечной памятью (распознающий автомат или *R*-автомат):

$$R_i^j = \langle X_i^j \times \hat{X}_i^{j+1}, 2^{Z_i^j}, X_i^{*j} \times \hat{X}_i^j, \varphi_i^j, \bar{\eta}_i^j, \rangle, \quad (1)$$

где (см. рис. 2)

- X_i^j — множество входных сигналов,
- X_i^{*j} — множество выходных сигналов,
- \hat{X}_i^{j+1} — множество управляющих сигналов с верхнего уровня иерархии,
- \hat{X}_i^j — множество управляющих сигналов на нижний уровень иерархии,
- $2^{Z_i^j}$ — множество состояний (множество подмножеств множества матриц предсказания),

- $\varphi_i^j : X_i^j \times \hat{X}_i^{j+1} \rightarrow 2^{\mathcal{Z}_i^j}$ — функция переходов,
- $\bar{\eta}_i^j : 2^{\mathcal{Z}_i^j} \rightarrow X_i^{*j} \times \hat{X}_i^j$ — вектор—функция выходов.

Множество входных признаков распознающего автомата R_i^j будем обозначать F_i^j , множество выходов — F_i^{*j} .

В качестве функции распознавания k -ого выходного признака \hat{f}_k в R -автомате удобно использовать набор булевых матриц предсказания $Z_k = \{Z_1^k, Z_2^k, \dots, Z_m^k\}$, в которых каждый столбец \bar{z}_u^r является вектором предсказания входных признаков в момент времени $\tau_s + u$, где τ_s — начало вычислительного цикла (момент действия управляющего сигнала \bar{x}_i^{j+1}). Сама матрица Z_r^k задаёт последовательность битовых векторов, наличие бита в котором свидетельствует о присутствии распознаваемого функцией \hat{f}_k признака. Алгоритм \mathcal{A}_{th} вычисления функции переходов φ_i^j и выходной функции $\bar{\eta}_i^j$ по начальному моменту времени τ_s , управляющему воздействию $\hat{x}_i^{j+1}(\tau_s)$ и входному воздействию ω_i^j представлен ниже.

1 Алгоритм \mathcal{A}_{th}

Вход: $\tau_s, \hat{x}_i^{j+1}(\tau_s), \omega_i^j$;

Выход: $\varphi_{i\Delta t}^j, \bar{\eta}_{i\Delta t}^j$;

- 1: $\hat{F}^* = \emptyset, Z^* = \emptyset, t = 0$; // активные функции распознавания и матрицы предсказания
 - 2: $c_1 \in (0, 1), c_2 \in (0, 1)$; // пороговые константы
// определение начального состояния
 - 3: для всех компонент \hat{x}_{ik}^{j+1} вектора $\hat{x}_i^{j+1}(\tau_s) = (\hat{x}_{i1}^{j+1}, \hat{x}_{i2}^{j+1}, \dots, \hat{x}_{il}^{j+1})$
 - 4: если $\hat{x}_{ik}^{j+1} \geq c_1$ то
 - 5: $\hat{F}^* := \hat{F}^* \cup \{\hat{f}_k\}$;
 - 6: $\bar{x}_i^j := \omega_i^j(\tau_s)$;
 - 7: для всех функций распознавания $\hat{f}_k \in \hat{F}^*$
 - 8: для всех $Z_r^k \in Z_k$, соответствующих функции распознавания \hat{f}_k ,
 - 9: если $\frac{\|\bar{z}_1^r - \bar{x}_i^j\|}{\|\bar{z}_1^r\| + \|\bar{x}_i^j\|} < c_2$ то
 - 10: $Z^* := Z^* \cup \{Z_r^k\}$;
 - 11: $\varphi_i^j(\bar{x}_i^j, \hat{x}_i^{j+1}(\tau_s)) := Z^*$; // значение функции переходов в начальный момент времени
 - 12: $\bar{N} := (|\{Z_r^1 | Z_r^1 \in Z^*\}|, \dots, |\{Z_r^{l_j} | Z_r^{l_j} \in Z^*\}|)$;
 - 13: $\eta(Z^*) = \bar{x}_i^{*j} := W(\bar{N})$; // значение функции выходов в начальный момент времени
 - 14: $\hat{x}_i^j = W(\sum_{\hat{f}_k \in \hat{F}^*} \hat{x}_{ik}^{j+1} \sum_{Z_r^k \in Z^*} \bar{z}_r^k)$;
// оновной цикл
 - 15: $t = 1$;
 - 16: пока $t \leq h_i^j - 1$
 - 17: $\bar{x}_i^j := \omega(\tau_s + t)$;
 - 18: для всех матриц предсказания Z_r^k из множества Z^*
 - 19: если $\frac{\|\bar{z}_{t+1}^r - \bar{x}_i^j\|}{\|\bar{z}_{t+1}^r\| + \|\bar{x}_i^j\|} \geq c_2$ то
 - 20: $Z^* := Z^* \setminus \{Z_r^k\}$;
 - 21: $\varphi_i^j(\bar{x}_i^j, \hat{x}_i^{j+1}(\tau_s)) := Z^*$; // значение функции переходов в момент времени t
 - 22: $\bar{N} = (|\{Z_r^1 | Z_r^1 \in Z^*\}|, \dots, |\{Z_r^{l_j} | Z_r^{l_j} \in Z^*\}|)$;
 - 23: $\eta(Z^*) = \bar{x}_i^{*j} := W(\bar{N})$; // значение функции выходов в момент времени t
 - 24: $t = t + 1$;
 - 25: если $t \leq h_i^j - 2$ то
 - 26: $\hat{x}_i^j := W(\sum_{\hat{f}_k \in \hat{F}^*} \hat{x}_{ik}^{j+1} \sum_{Z_r^k \in Z^*} \bar{z}_r^k)$;
-

В алгоритме используется стандартная функция W нормировки весовых значений:

$$W(\bar{x}) = \left(\frac{x_1}{\max_i x_i}, \dots, \frac{x_n}{\max_i x_i} \right), \quad (2)$$

где $\bar{x} = (x_1, \dots, x_n)$ — вектор с ненормированными компонентами.

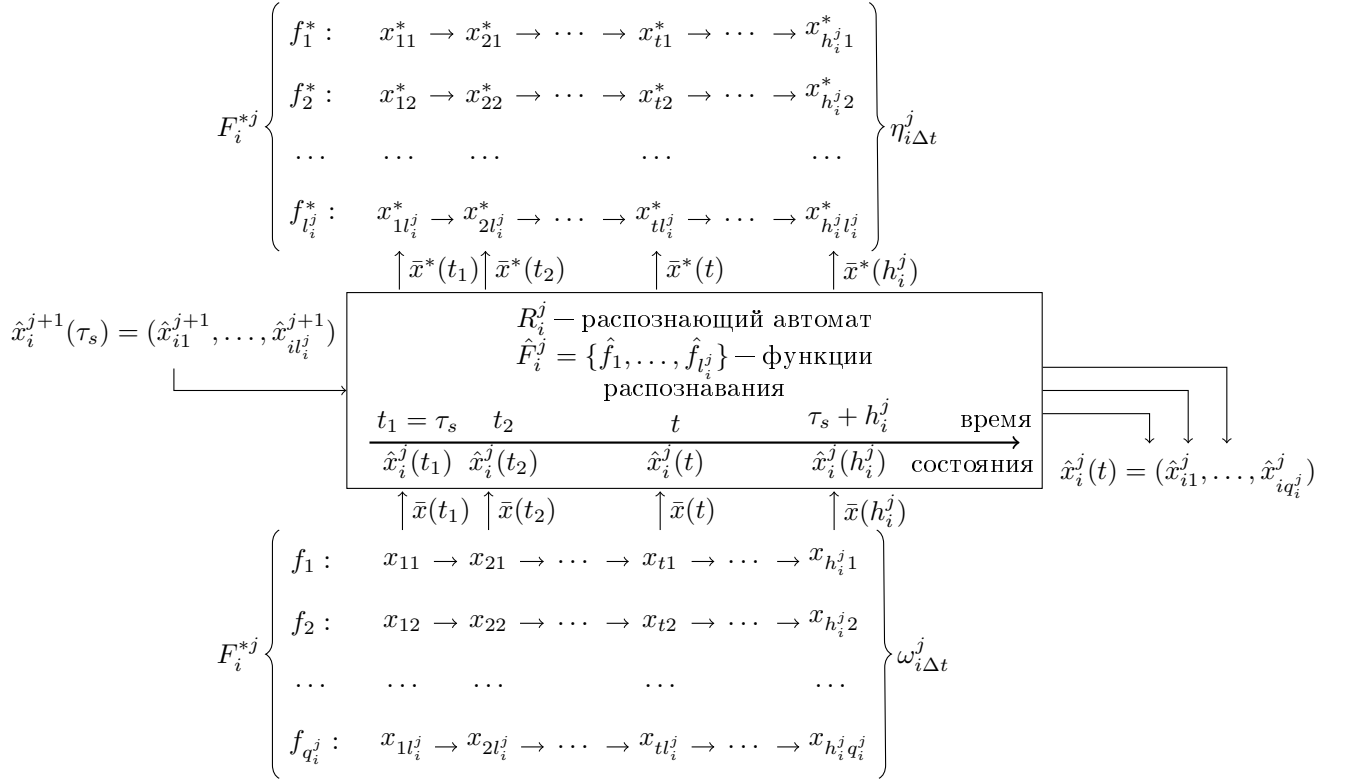


Рис. 2: Входные и выходные сигналы распознающего автомата.

1.1. Определение компонент знака. Для определения компонент знака через описанный в предыдущем разделе R -автомат необходимо ввести ряд вспомогательных понятий.

Введём семейство бинарных отношений $\{\sqsubset, \sqsubset^1, \sqsubset^2, \dots\}$, определённых на декартовом произведении $\{f_k\} \times \{f_k\}$. Будем считать, что «признак f_1 является составляющим признака f_2 » или «признак f_2 измеряется по признаку f_1 », $(f_1, f_2) \in \sqsubset$ или $f_1 \sqsubset f_2$, в том случае, если $f_1 \dashv R_1^j$, $f_2 \dashv R_2^{j+1}$, R_2^{j+1} — родительский блок по отношению к R_1^j и в множестве матриц предсказания \mathcal{Z}_2 признака f_2 существует как минимум одна матрица Z_r^2 , содержащая некоторый столбец \bar{z}_u^r с элементом $z_{uv}^r \neq 0$, где v — индекс признака f_1 во входном векторе для распознающего блока R_2^{j+1} .

Пара признаков $(f_1, f_2) \in \sqsubset^t$ или $f_1 \sqsubset^t f_2$, где $t \in \{1, 2, \dots\}$, в том случае, если $f_1 \dashv R_1^j$, $f_2 \dashv R_2^{j+1}$, R_2^{j+1} — родительский блок по отношению к R_1^j и в множестве матриц предсказания \mathcal{Z}_2 признака f_2 существует как минимум одна матрица Z_r^2 , содержащая t -ый столбец \bar{z}_t^r с элементом $z_{tv}^r \neq 0$, где v — индекс признака f_1 во входном векторе для распознающего блока R_2^{j+1} .

Каждый элемент векторов—столбцов соотносится с признаком из входного множества признаков распознающего блока, что означает задание типа для каждого элемента вектора—столбца. Будем обозначать тип k -го элемента вектора-столбца распознающего блока R_i^j как $f_i^j(k) \in F_i^j$, $k \in (1, q_i^j)$.

Введём два выделенных из множества $\{f_k\}$ признака: f_c — «условие» и f_e — «эффект», распознаваемые одним распознающим блоком R_0^1 : $F_0^{*1} = \{f_c, f_e\}$. Те признаки, которые распознаются распознающими блоками, выступающими родительскими по отношению к блоку R_0^1 , будем называть процедурными признаками, остальные — объектными признаками.

Для любого процедурного признака выполняются следующие естественные условия:

- условие всегда предшествует эффекту,
- условие всегда влечёт за собой эффект и
- все условия всегда отделены от своих эффектов.

Иными словами, если f_1 — процедурный признак, то если в столбце \bar{z}_u^r матрицы предсказания Z_r^1 элемент z_{uv}^r , соответствующий признаку f_c , не равен 0, то в этом столбце соответствующий признаку f_e элемент вектора — нулевой, в столбце $z_{u+t}^r, t > 0$ наоборот — элемент $z_{u+t,v}^r$, соответствующий признаку f_c , равен 0, а соответствующий признаку f_e элемент — не нулевой. Те столбцы матрицы предсказания Z , в которых соответствующий признаку f_e элемент вектора не нулевой, будем называть *столбцами эффектов*, а те столбцы матрицы предсказания Z , в которых не равен нулю элемент вектора, соответствующий признаку f_c — *столбцами условий*.

Пополним семейство отношений $\{\sqsubset, \sqsubset^1, \sqsubset^2, \dots\}$ двумя отношениями: \sqsubset^c и \sqsubset^e , принадлежность к которым пары признаков (f_1, f_2) свидетельствует о том, что признак f_1 присутствует соответственно в столбце условий и эффектов как минимум в одной матрице предсказания процедурного признака f_2 .

Определение 1. Если f_1 — признак, то подмножество $\tilde{p}(f_1)$ множества $\{f_k\}$ таких признаков, что $\forall f_i \in \tilde{p}(f_1) f_i \sqsubset f_1$, будем называть перцептом признака f_1 .

На множестве всех перцептов \tilde{P} введём величину $\rho_p(\tilde{p}(f_1), \tilde{p}(f_2))$, вычисляемую по следующему правилу:

- если f_1 и f_2 распознаются разными распознающими блоками, т. е. $f_1 \dashv R_1^j, f_2 \dashv R_2^i$, то $\rho_p(\tilde{p}(f_1), \tilde{p}(f_2)) = \infty$,
- если f_1 и f_2 распознаются одним и тем же распознающим блоком R_1^j со множеством входных признаков F_1^j мощности q и характерным временем h , то

$$\rho_p(\tilde{p}(f_1), \tilde{p}(f_2)) = \min_{\substack{Z_r^1 \in Z_1 \\ Z_s^2 \in Z_2}} \frac{1}{q \cdot h} \sum_{u=1}^h \|\bar{z}_u^r - \bar{z}_u^s\|. \quad (3)$$

Утверждение 1. Величина ρ_p является метрикой на множестве перцептов \tilde{P} .

Определение 2. Если f_1 — признак, f_2 — процедурный признак, $f_1 \sqsubset^c f_2$, то будем называть f_2 функциональным значением признака f_1 . Множество всех функциональных значений признака f_1 будем обозначать $\tilde{m}(f_1)$.

На множестве всех функциональных значений \tilde{M} введём величину $\rho_m(\tilde{m}(f_1), \tilde{m}(f_2))$, вычисляемую по следующему правилу:

$$\rho_m(\tilde{m}_1(f_1), \tilde{m}_2(f_2)) = \min_{\substack{f_i \in \tilde{m}(f_1) \\ f_j \in \tilde{m}(f_2)}} \rho_p(\tilde{p}(f_i), \tilde{p}(f_j)). \quad (4)$$

Утверждение 2. Величина ρ_m является метрикой на множестве функциональных значений \tilde{M} .

1.2. Семантический уровень обобщения. Определение ролевой структуры для алгоритма планирования.

Здесь я бы предложил дать определение типам обобщения на семантическом уровне.

2.Связывание образа и значения. Для формальной записи итерационного процесса связывания образа и значения в алгоритме образования знака из первой части статьи рассмотрим подробнее строение матрицы предсказания процедурного признака. Матрицу предсказания Z_r^p процедурного признака f_p всегда можно представить в следующем виде:

$$Z_r^p = (\bar{z}_1^{r,c}, \dots, \bar{z}_{j_1}^{r,c}, \bar{z}_{j_1+1}^{r,e}, \dots, \bar{z}_{i_1}^{r,e}, \dots, \bar{z}_{i_{k-1}+1}^{r,c}, \dots, \bar{z}_{j_k}^{r,c}, \bar{z}_{j_k+1}^{r,e}, \dots, \bar{z}_{i_k}^{r,e}), \quad (5)$$

где $\bar{z}_j^{r,c}$ — столбцы причин, $\bar{z}_i^{r,e}$ — столбцы следствий.

Величину k будем называть актностью процедурного признака. В дальнейшем будем рассматривать простые матрицы предсказаний k -актного процедурного признака:

$$Z_r^p = (\bar{z}_1^{r,c}, \bar{z}_2^{r,e}, \dots, \bar{z}_{2 \cdot k-1}^{r,c}, \bar{z}_{2 \cdot k}^{r,e}). \quad (6)$$

Краткая форма k -актного процедурного признака f_p имеет матрицу предсказания, в которой оставлены только первый столбец условий и последний столбец эффектов.

Любой одноактный процедурный признак f_p , распознаваемый блоком R_i^j , можно представить в виде правила $r_p = (F_C(f_p), F_A(f_p), F_D(f_p))$, в котором:

- $F_C(f_p) \subseteq F_i^j$ — множество признаков — условий правила: $\forall f \in F_C(f_p) f \sqsubset^c f_p$;
- $F_A(f_p) \subseteq F_i^j$ — множество добавляемых правилом признаков: $\forall f \in F_A(f_p) f \sqsubset^e f_p, f \notin F_C$;
- $F_D(f_p) \subseteq F_i^j$ — множество удаляемых правилом признаков: $\forall f \in F_D(f_p) f \notin F_A, f \in F_C$.

Очевидно, выполняются следующие соотношения: $F_A(f_p) \cap F_D(f_p) = \emptyset, F_A(f_p) \cap F_C(f_p) = \emptyset, F_D(f_p) \subseteq F_C(f_p)$.

Определение 3. *Процедурный признак f_p^1 с матрицей предсказания $Z = (\bar{z}_1^c, \bar{z}_2^e)$ выполняется на векторе z длины q , если $z \cdot \bar{z}_1^c = \bar{z}_1^c$.*

Будем говорить, что процедурный признак f_p^1 выполним в условиях процедурного признака f_p^2 , если

- оба признака распознаются одним и тем же распознающим блоком R_i^j и признак f_p^1 выполняется на столбце условий матрицы предсказания признака f_p^2 ,
- $f_p^1 \dashv R_1^{j_1}, f_p^2 \dashv R_2^{j_2}$, множества $F_C(f_p^1)$ и $F_C(f_p^2)$ состоят из одних и тех же признаков, образуемый вектор \tilde{z} (той же мощности, что и множество $F_1^{j_1}$) элементы которого, соответствующие признакам из $F_C(f_p^2)$ принимаются равными 1, остальные — 0, и признак f_p^1 выполним на векторе \tilde{z} .

Определение 4. *Будем говорить, что два процедурных признака f_p^1 и f_p^2 конфликтуют, если выполнено как минимум одно из следующих условий:*

- $F_D(f_p^1) \cap F_A(f_p^2) \neq \emptyset$,
- $F_D(f_p^2) \cap F_A(f_p^1) \neq \emptyset$,
- $F_D(f_p^1) \cap F_C(f_p^2) \neq \emptyset$,

- $F_D(f_p^2) \cap F_C(f_p^1) \neq \emptyset$.

Определение 5. Результатом операции сохраняющего приведения вектор—столбца \bar{z}_1 к множеству входных признаков $F_{i_2}^{j_2}$ будем называть такой вектор \bar{z}_3 длины $q_{i_2}^{j_2}$, элемент которого $z_{3k} = 1$, если $f_{i_1}^{j_1}(k) = f_{i_2}^{j_2}(k)$ и $z_{1k} = 1$, иначе $z_{3k} = 0$, и обозначать $(\bar{z}_1 \rightarrow F_{i_2}^{j_2}) = \bar{z}_3$.

Определение 6. Результатом операции сужающего приведения вектор—столбца \bar{z}_1 к некоторому столбцу \bar{z}_2 распознающего блока $R_{i_2}^{j_2}$ будем называть такой вектор \bar{z}_3 длины $q_{i_2}^{j_2}$, элемент которого $z_{3k} = 1$, если $f_{i_1}^{j_1}(k) = f_{i_2}^{j_2}(k)$, $z_{2k} = 1$ и $z_{1k} = 1$, иначе $z_{3k} = 0$, и обозначать $(\bar{z}_1 \Rightarrow \bar{z}_2) = \bar{z}_3$.

Будем считать, что у субъекта имеется опыт наблюдения, который выражается в виде отношения $\Psi_p^m: \tilde{p}\Psi_p^m\tilde{m}$, или $\Psi_p^m(\tilde{p}) = \tilde{m}$, в том случае, если $\tilde{p} \in \tilde{P}$ является перцептом некоторого признака f , а $\tilde{m} \in \tilde{M}$ — функциональным значением того же признака f .

Ниже представлен алгоритм доопределения функции Ψ_p^m , который и отражает собой суть итерационного процесса во время образования знака согласно алгоритму из первой части статьи. Доопределение проводится на новую пару (\tilde{p}, \tilde{m}) , где функциональное значение \tilde{m} строится в сравнении с эталоном \tilde{m}^0 , а перцепт \tilde{p} формируется на основе подмножества составляющих признаков \hat{F} . Доопределение функции Ψ_p^m означает формирование нового признака f^* , т. е. его первой матрицы предсказания Z^* в рамках распознающего блока R^* .

2 Алгоритм \mathfrak{A}_{pm}

Вход: $\tilde{m}^0 = \{f_p\}$, $\Psi_p^m, \hat{F} \subseteq \{f_k\}$;

- 1: $\tilde{p}^{*(0)} := \emptyset$;
- 2: $Z^{*(0)} := \emptyset$;
- 3: $t := 0$;
- 4: **для всех** $f^{(t)} \in \hat{F}$
- 5: **если** $\exists \tilde{m}^{(t)} \in \tilde{M}$ такое, что $(\tilde{p}(f^{(t)}), \tilde{m}^{(t)}) \in \Psi_p^m$ **and** $\tilde{m}^{(t)}$ выполним в условиях признака f_p **and** $\nexists f: f \in \tilde{p}^{*(t)}, (\tilde{p}(f), \tilde{m}(f)) \in \Psi_p^m, \tilde{m}^0$ конфликтует с $\tilde{m}^{(t)}$ **то**
- 6: $\tilde{p}^{*(t)} = \tilde{p}^{*(t)} \cup \{f^{(t)}\}$;
- 7: **если** $\exists R_i^j$ такой, что $f^{(t)} \in F_i^j$ **то**
- 8: $R_i^{j(t)} := R_i^j$;
- 9: **иначе**
- 10: $R_i^{j(t)} := \arg \max_{\{R\}} (F_i^j \cap \tilde{p}^{(t)}), F_i^{j(t)} := F_i^{j(t)} \cup f^{(t)}$;
- 11: $\bar{z}_s := (z_{s1}, z_{s2}, \dots, z_{sq}), z_{sk} = 1$, если k — индекс признака $f^{(t)}$ во входном векторе распознающего блока $R_i^{j(t)}$ и $z_{sk} = 0$ иначе;
- 12: $Z^{*(t)} := Z^{*(t)} \cup \bar{z}_s$;
- 13: $Z_p^{(t)} := (\bar{z}_1^{c(t)}, \bar{z}_2^{e(t)}, \dots, \bar{z}_{2 \cdot k-1}^{c(t)}, \bar{z}_{2 \cdot k}^{e(t)})$, где $\bar{z}_i^{c(t)} = \bigvee_{\tilde{m}_j^{(t)}} (\bar{z}_j^{c(t)} \rightarrow F_p^j)$,
- 14: $\bar{z}_i^{e(t)} = \bigvee_{\tilde{m}_j^{(t)}} (\bar{z}_j^{e(t)} \Rightarrow \bar{z}_j^e)$;
- 15: $\tilde{m}^{*(t)} = \{f_p^{(t)}\}$;
- 16: $Z^{*(t)} = \{Z^{*(t)}\}$;
- 17: $t = t + 1$;

вернуть Ψ_p^m , определённая на паре (\tilde{p}, \tilde{m}) , где $\tilde{p} = \lim_{t \rightarrow |\hat{F}|} \tilde{p}^{*(t)}$, $\tilde{m} = \lim_{t \rightarrow |\hat{F}|} \tilde{m}^{*(t)}$, $f^*, Z^* =$

$\lim_{t \rightarrow |\hat{F}|} Z^{*(t)}, Z^* = \{Z^*\}$;

Теорема 1 (о корректности алгоритма \mathfrak{A}_{pm}). Алгоритм \mathfrak{A}_{pm} корректен, т. е. последовательность функциональных значений $\langle \tilde{m}^{*(0)}, \tilde{m}^{*(1)}, \dots \rangle$, которая строится с помощью алгоритма \mathfrak{A}_{pm} для функционального значения \tilde{m}^0 , сходится к \tilde{m}^0 .

Доказательство. Рассмотрим два элемента последовательности $\tilde{m}^{*(t)} = \{f_p^{(t)}\}$ и $\tilde{m}^{*(t+1)} = \{f_p^{(t+1)}\}$. Соответствующие матрицы предсказания будут иметь следующий вид:

$$Z_p^{(t)} = (\bar{z}_1^{c(t)}, \bar{z}_2^{e(t)}, \dots, \bar{z}_{2 \cdot k - 1}^{c(t)}, \bar{z}_{2 \cdot k}^{e(t)}), \quad (7)$$

$$Z_p^{(t+1)} = (\bar{z}_1^{c(t+1)}, \bar{z}_2^{e(t+1)}, \dots, \bar{z}_{2 \cdot k - 1}^{c(t+1)}, \bar{z}_{2 \cdot k}^{e(t+1)}). \quad (8)$$

Если на шаге 1 и 2 алгоритма \mathfrak{A}_{pm} на $(t+1)$ -й итерации не был найден подходящий признак, то матрицы $Z_p^{(t)}$ и $Z_p^{(t+1)}$ равны. Рассмотрим случай, когда был найден подходящий признак $f^{(t+1)}$ с функциональным значением $\tilde{m}^{(t+1)} = \{\tilde{f}_p^{(t+1)}\}$ с соответствующей матрицей предсказания $\tilde{Z}_p^{(t+1)} = (\bar{z}^{c(t+1)}, \bar{z}^{e(t+1)})$.

Т. к. выполнено условие шага 1, то признак $\tilde{f}_p^{(t+1)}$ выполним на некотором $(2 \cdot s - 1)$ -м столбце условий матрицы предсказания признака f_p . Это означает, что матрицы $Z_p^{(t)}$ и $Z_p^{(t+1)}$ будут отличаться только в двух вектор-столбцах $(2 \cdot s - 1)$ -м и $(2 \cdot s)$ -м:

$$\bar{z}_{2 \cdot s - 1}^{c(t+1)} = \bar{z}_{2 \cdot s - 1}^{c(t)} \vee (\bar{z}^{c(t+1)} \rightarrow F_p^j), \bar{z}_{2 \cdot s}^{e(t+1)} = \bar{z}_{2 \cdot s}^{e(t)} \vee (\bar{z}^{e(t+1)} \Rightarrow \bar{z}_{2 \cdot s}^e). \quad (9)$$

По определению расстояние между функциональными значениями $\tilde{m}^{(t)}$ и \tilde{m}^0 примет следующее значение:

$$\begin{aligned} \rho_m(\tilde{m}^{(t)}, \tilde{m}^0) &= \min_{\substack{f_i \in \tilde{m}^{(t)} \\ f_j \in \tilde{m}^0}} \rho_p(\tilde{p}(f_i), \tilde{p}(f_j)) = \rho_p(\tilde{p}(f_p^{(t)}), \tilde{p}(f_p)) = \\ &= \frac{1}{q \cdot h} \sum_{\substack{\bar{z}_u^1 \in Z_p^{(t)} \\ \bar{z}_u^2 \in Z_p}} \|\bar{z}_u^1 - \bar{z}_u^2\|. \end{aligned} \quad (10)$$

Аналогично для $\tilde{m}^{(t+1)}$:

$$\rho_m(\tilde{m}^{(t+1)}, \tilde{m}^0) = \frac{1}{q \cdot h} \sum_{\substack{\bar{z}_u^1 \in Z_p^{(t+1)} \\ \bar{z}_u^2 \in Z_p}} \|\bar{z}_u^1 - \bar{z}_u^2\|. \quad (11)$$

Рассмотрим разность

$$\begin{aligned} \rho_m(\tilde{m}^{(t)}, \tilde{m}^0) - \rho_m(\tilde{m}^{(t+1)}, \tilde{m}^0) &= \frac{1}{q \cdot h} (\|\bar{z}_{2 \cdot s - 1}^{c(t)} - \bar{z}_{2 \cdot s - 1}^c\| + \|\bar{z}_{2 \cdot s}^{e(t)} - \bar{z}_{2 \cdot s}^e\| - \\ &- \|\bar{z}_{2 \cdot s - 1}^{c(t+1)} - \bar{z}_{2 \cdot s - 1}^c\| - \|\bar{z}_{2 \cdot s}^{e(t+1)} - \bar{z}_{2 \cdot s}^e\|) = \frac{1}{q \cdot h} (\|\bar{z}_{2 \cdot s - 1}^{c(t)} - \bar{z}_{2 \cdot s - 1}^c\| + \\ &+ \|\bar{z}_{2 \cdot s}^{e(t)} - \bar{z}_{2 \cdot s}^e\| - \|\bar{z}_{2 \cdot s - 1}^{c(t)} \vee (\bar{z}^{c(t+1)} \rightarrow F_p^j) - \bar{z}_{2 \cdot s - 1}^c\| - \\ &- \|\bar{z}_{2 \cdot s}^{e(t)} \vee (\bar{z}^{e(t+1)} \Rightarrow \bar{z}_{2 \cdot s}^e) - \bar{z}_{2 \cdot s}^e\|), \end{aligned} \quad (12)$$

где $\bar{z}_{2 \cdot s - 1}^c, \bar{z}_{2 \cdot s}^e$ — столбцы матрицы предсказания процедурного признака f_p , соответствующего функциональному значению \tilde{m}^0 .

Так как $\tilde{f}_p^{(t+1)}$ выполним на $(2 \cdot s - 1)$ -м столбце условий матрицы предсказания признака f_p , то после применения операции приведения $\bar{z}^{c(t+1)} \rightarrow F_p^j$ в результирующем векторе единицы появляются только на тех же местах что и в векторе $\bar{z}_{2 \cdot s - 1}^c$.

Это означает, что в векторе $\bar{z}_{2.s-1}^{c(t)} \vee (\bar{z}^{c(t+1)} \rightarrow F_p^j)$ по сравнению с вектором $\bar{z}_{2.s-1}^{c(t)}$ единицы находятся только в тех же местах, что и в векторе $\bar{z}_{2.s-1}^c$, а новых нулей не появляется. В следствие чего разность $\|\bar{z}_{2.s-1}^{c(t)} - \bar{z}_{2.s-1}^c\| - \|\bar{z}_{2.s-1}^{c(t)} \vee (\bar{z}^{c(t+1)} \rightarrow F_p^j) - \bar{z}_{2.s-1}^c\|$ всегда больше нуля.

Так как для столбцов эффектов применяется операция сужающего приведения, которая оставляет единицы только на тех местах, на которых одновременно находятся единицы в приводимом векторе и векторе, к которому осуществляется приведение. В связи с этим разность $\|\bar{z}_{2.s}^{e(t)} - \bar{z}_{2.s}^e\| - \|\bar{z}_{2.s}^{e(t)} \vee (\bar{z}^{e(t+1)} \Rightarrow \bar{z}_{2.s}^e) - \bar{z}_{2.s}^e\|$ также больше нуля.

Так как обе разности в скобках выражения для $\rho_m(\tilde{m}^{(t)}, \tilde{m}^0) - \rho_m(\tilde{m}^{(t+1)}, \tilde{m}^0)$ больше нуля, то отсюда следует, что функциональное значение $\tilde{m}^{(t+1)}$ ближе к \tilde{m}^0 . В виду произвольности выбора итерации t , это приводит к сходимости всей последовательности $\langle \tilde{m}^{*(0)}, \tilde{m}^{*(1)}, \dots \rangle$. ■

3. Самосознание и его функции. Для алгоритма планирования.

Здесь я бы предложил дать психологическое описание функций самосознания и определения функций оценки Φ_a и Φ_p .

4. Алгоритм планирования. Планом *Plan* будем называть такую последовательность личностных смыслов, в которой действие, представляемое очередным личностным смыслом не конфликтует с предыдущим в цепочке действием.

Целевая ситуация строится исходя из образа процедурного признака, связанного с личностным смыслом, который был определён в процессе целеполагания для целевого знака (см. первую часть статьи).

На странице 10 представлен алгоритм планирования поведения.

Заключение. Подведение общих итогов.

Про архитектуру агентов и распределение ролей.

СПИСОК ЛИТЕРАТУРЫ

1. Осипов Г. С., Панов А. И., Чудова Н. В. Управление поведением как функция сознания. I. Картина мира и целеполагание // Известия РАН. Теория и системы управления. — 2014. — № 4. — С. 83–96.
2. Лурия А. Р. Мозг и психические процессы. Т. 2. — М. : Педагогика, 1970.
3. Выготский Л. С. Психология развития человека. — М. : Издательство Смысл, 2005. — С. 1136.
4. Леонтьев А. Н. Деятельность. Сознание. Личность. — М. : Политиздат, 1975.
5. Артемьева Е. Ю. Психология субъективной семантики. — М. : Издательство МГУ, 1980.
6. Эделмен Д., Маунткасл В. Разумный мозг. — М. : Мир, 1981.
7. Иваницкий А. М. Мозговая основа субъективных переживаний: гипотеза информационного синтеза // Журнал высшей нервной деятельности. — 1996. — Т. 46, № 2. — С. 241–282.

3 Алгоритм \mathcal{A}_{bp}

Вход: начальная ситуация F_{sit} , целевая ситуация F_{goal} , функции оценки Φ_a и Φ_p ;

Выход: план $Plan$;

```
1:  $Plan = \text{PLANNING}(\emptyset, F_{goal});$ 
2: процедура  $\text{PLANNING}(Plan, F_{cur})$ 
3:    $\Delta = F_{sit} \setminus F_{cur};$  // текущая невязка состояний
4:    $F_{for} = \arg \min_{F \in 2^{F_{sit}}} |\bigcap_{f_p \in F} F_A(f_p) \setminus \Delta|;$  // находим множество наиболее подходящих парал-
      лельных действий
5:   для всех  $f_j \in F_{for}$ 
6:     если  $\exists f_k \in F_{for}$  такой, что  $f_k \neq f_j$  и  $f_k$  конфликтует с  $f_j$  то
7:        $F_{for} = F_{for} \setminus \{f_k\};$  // Удаляем конфликтующие признаки
8:    $F_a^{for} = \emptyset;$  // текущее множество личностных смыслов
9:   для всех  $f_p \in F_{for}$ 
10:     $F_a^{for} = F_a^{for} \cup \{\text{INTERIOR}(f_p)\};$  // интериоризация значения
11:   $\tilde{F}_a^{for} = \Phi_a(F_a^{for}, f_{goal});$  // выбор предпочитаемых действий
12:  если  $\bigcup_{f \in \tilde{F}_a^{for}} F_C(f) \subseteq F_{sit}$  то
13:    вернуть  $Plan \cup \tilde{F}_a^{for};$  // возвращаем обновленный план
14:  иначе
15:     $\Delta^* = \Phi_p(\Delta, f_{goal});$  // Ранжирование критических признаков
16:     $\tilde{F}_a^{back} = \emptyset;$ 
17:    для всех  $f_k \in \Delta^*$ 
18:       $m_k = \tilde{m}(f_k);$  // определение значение  $k$ -го знака
19:       $F_a^{back} = \emptyset;$ 
20:      для всех  $f_p \in m_k$ 
21:         $F_a^{back} = F_a^{back} \cup \{\text{INTERIOR}(f_p)\};$ 
22:       $\tilde{F}_a^{back} = \tilde{F}_a^{back} \cup \Phi_a(F_a^{back}, f_{goal});$  // выбор предпочитаемых действий
23:    для всех  $f_j \in \tilde{F}_a^{back}$ 
24:      если  $\exists f_k \in \tilde{F}_a^{back}$  такой, что  $f_k \neq f_j$  и  $f_k$  конфликтует с  $f_j$  то
25:         $\tilde{F}_a^{back} = \tilde{F}_a^{back} \setminus \{f_k\};$  // Удаляем конфликтующие признаки
26:    если  $\Delta \not\subseteq \bigcup_{f \in \tilde{F}_a^{back}} F_A(f)$  то
27:      вернуть невозможно построить план;
28:    иначе
29:      вернуть  $\text{PLANNING}(Plan, \bigcup_{f \in \tilde{F}_a^{back}} F_C(f));$ 
```
