УДК 519.21;519.714.3;681.518 DOI: http://doi.org/10.25728/pu.2020.2.4

СОГЛАСОВАННОЕ УПРАВЛЕНИЕ МНОГОЭЛЕМЕНТНЫМИ ДИНАМИЧЕСКИМИ ОРГАНИЗАЦИОННЫМИ СИСТЕМАМИ. Ч. 2. Многоуровневая динамическая организационная система

М.В. Белов

Аннотация. Полученные в первой части статьи результаты распространены на многоуровневую динамическую многоэлементную организационную/активную систему (ОС), а также на случай неопределенных затрат агентов. Доказаны утверждения о том, что для любой допустимой траектории результатов может быть построена согласованная компенсаторная система стимулирования, которая реализует (как равновесие в доминантных стратегиях) траекторию действий агентов, приводящих к требуемой траектории результатов; декомпозирует задачу управления по агентам и по периодам времени; гарантированно обеспечивает (по всем возможным дальновидностям агентов) минимальные затраты управляющего органа (центра) на реализацию данной траектории результатов. Показано, что в таких системах стимулирования размеры платежей зависят только от соответствующих значений функций затрат, которые, в свою очередь, косвенно учитывают технологические функции, структуру сетей и структуру ОС в целом. Поставлена задача оптимального планирования и указан алгоритм ее решения.

Ключевые слова: стимулирование, многоуровневые динамические активные системы, согласованное управление.

ВВЕДЕНИЕ

В данной части статьи продолжается исследование динамических организационных систем, в которых агенты реализуют свои действия согласно сложным сетевым технологиям. Такие технологии задают взаимозависимости между действиями агентов и их результатами, что в существенной степени отвечает практике современных фирм, корпораций и предприятий других форм. В первой части [1] статьи исследована динамическая организационная система в составе одного центра и множества агентов. В настоящей работе постановка задачи согласованного управления динамическими ОС усложняется — вводится иерархия уровней в структуре динамической ОС (§ 3), строится компенсаторная система стимулирования, декомпозирующая игру по уровням, агентам и периодам времени (§ 4). Наконец, в § 5 рассмотрены принципы построения алгоритмов оптимального согласованного планирования в многоэлементных динамических ОС[1].

3. МНОГОУРОВНЕВЫЕ ДИНАМИЧЕСКИЕ ОРГАНИЗАЦИОННЫЕ СИСТЕМЫ

Основываясь на результатах, полученных в первой части статьи [1], усложним модель организационной системы — введем иерархию подсистем, каждая из которых является динамической сете-

[1] Нумерация параграфов, формул и утверждений продолжает нумерацию первой части [1] статьи.

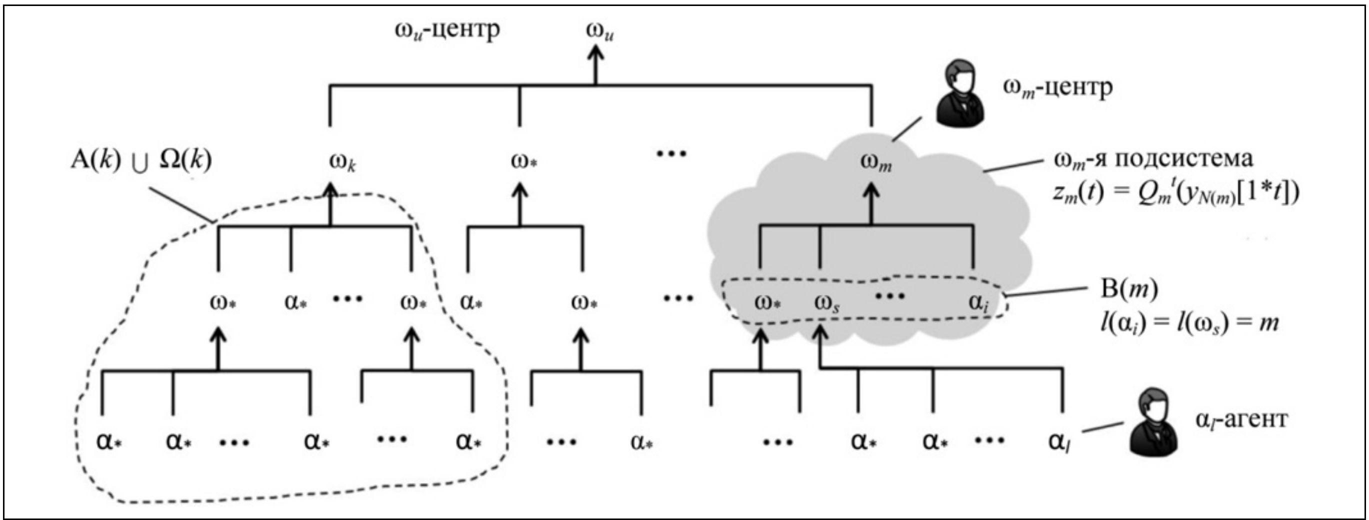

Рис. 1. Иерархия активных элементов — агентов и промежуточных центров

вой организационной системой (ДСОС). Сокращенно будем называть такую систему многоуровневой ДСОС, м-ДСОС (подразумевая, что каждая из подсистем образует ДСОС)[2]. Иерархическую структуру такой многоуровневой ОС (рис. 1) зададим иерархией *активных элементов* (АЭ).

Некоторые из АЭ, обозначив их множество как $A = \{\alpha_1, ..., \alpha_a\}$, будем называть агентами и применять к конкретному АЭ обозначение α_i-агент. Активные элементы, выступающие в роли центров подсистем[3] ОС (ω_m-подсистем, их множество обозначим $\Omega = \{\omega_1, ..., \omega_u\}$), будем называть, следуя работе [2], *промежуточными центрами*, п-центрами или ω_m-центрами.

Иерархией АЭ назовем связный ациклический граф $\Gamma = \langle A \cup \Omega, D\rangle$ (где $A \cup \Omega$ — множество вершин, а $D \subseteq (A \cup \Omega) \times \Omega$ — множество дуг $d_{im} \in E$), такой что:

— ни для одной из вершин $\alpha_i \in A$ не существует входящих дуг ε_{li},

— существует единственная вершина, *корневая вершина* (считаем, что это вершина ω_u), в которую существуют пути из всех вершин $\alpha_i \in A$ и $\omega_m \in \Omega$ ($\omega_m \neq \omega_u$),

— из каждой вершины (кроме вершины ω_u) выходит ровно одна дуга.

Для каждой вершины $\omega_m \in \Omega$ определим $A(m) \subseteq A$ — подмножество вершин-агентов, из которых существует путь в вершину ω_m, аналогично $\Omega(m) \subseteq \Omega$ — подмножество промежуточных центров, из которых существует путь в вершину ω_m. Тогда $A(m) \cup \Omega(m)$ образует множество подсистем и агентов, нижестоящих относительно ω_m-подсистемы. Каждой из вершин графа Г также соответствует один и только один АЭ: вершинам α_i — агенты, вершинам ω_m — п-центры, вершине ω_u — центр верхнего уровня.

Корневая вершина ω_u соответствует цели и результату деятельности ОС в целом.

На множестве всех АЭ $(A \cup \Omega)$ зададим единую нумерацию элементов, назовем ее технологической или т-нумерацией. В дальнейшем в качестве индексов ω_i или α_i будем употреблять т-номера так, что ω_i или α_i однозначно идентифицируют п-центр и/или агента. Т-нумерацию зададим алгоритмом, учитывающим, что на иерархии Г, в силу ее свойств, может быть задана правильная нумерация [3] всех АЭ, а в каждой подсистеме также может быть задана правильная нумерация на множестве АЭ соответственно технологической сети G_m.

Шаг 1. Согласно иерархии Г зададим промежуточную правильную нумерацию на множестве всех АЭ (агентов и п-центров), будем называть ее номера Г-номерами, таким образом обозначая его отличие от технологической нумерации. Введем вспомогательную целую переменную η и присво-

[2] Иерархия сетей с правильной нумерацией может быть редуцирована в единственную сеть с правильной нумерацией, и в соответствии с технологией выполнения агентами действий и получения результата м-ДСОС сводится к ДСОС. Однако над такой технологической сетью построена иерархия стимулирующих (управляющих) промежуточных центров, что отличает данную постановку от рассмотренной в § 1 и 2 [1].

[3] Важно отметить, что подсистемы ОС являются ОС, что отражает фрактальный характер данной модели.

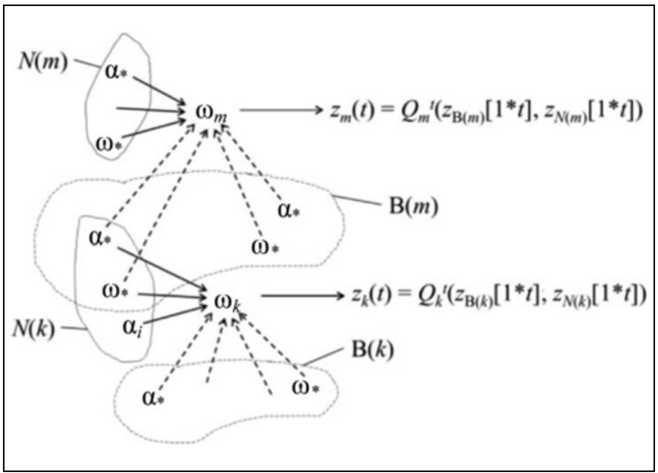

Рис. 2. Технологические связи в подсистемах многоуровневой динамической сетевой организационной системы

им $\eta = 0$. Считаем, что ни одному АЭ не назначен т-номер.

Шаг 2. Среди всех п-центров, которым еще не назначен т-номер, выбираем тот, у которого Г-номер минимален, пусть таким оказался ω_μ-центр. Для всех АЭ (агентов и п-центров), принадлежащих ω_μ-подсистеме, выполняем правильную нумерацию согласно сети G_μ, соответствующие номера будем называть G-номерами.

Шаг 3. Полагаем $\eta = \eta + 1$. Среди всех агентов, принадлежащих ω_μ-подсистеме, которым еще не назначен т-номер, выбираем тот, у которого Г-номер минимален, и присваиваем ему т-номер, равный η.

Шаг 4. Проверяем, есть ли еще агенты, принадлежащие ω_μ-подсистеме, которым не назначен т-номер. Если такие агенты есть, возвращаемся к шагу 3, иначе переходим к шагу 5.

Шаг 5. Полагаем $\eta = \eta + 1$. Присваиваем ω_μ-центру т-номер, равный η.

Шаг 6. Проверяем, есть ли еще п-центры, которым не назначен т-номер. Если такие п-центры есть, возвращаемся к шагу 2, иначе алгоритм завершен.

Полученная т-нумерация обладает важным свойством: результат действий любого из АЭ зависит от действий самого АЭ и результатов АЭ с меньшими т-номерами и не зависит от действий АЭ с бóльшими т-номерами. В дальнейшем, если не оговорено особо, будем считать, что при ссылке на номер АЭ имеется в виду технологический номер, префикс т- при этом указываться не будет. В определенном смысле с помощью «перенумерации» элементов ОС удается свести «иерархию сетей» к одной эквивалентной сети с правильной нумерацией.

Обозначим В(m) множество АЭ в ω_m-й подсистеме ($\mathrm{B}(m) \subseteq \mathrm{A}(m) \cup \Omega(m)$, см. рис. 1): α_i-агентов таких, что для $\forall \alpha_i \in \mathrm{B}(m)\ \exists\ d_{im} \in D$, и ω_k-центров таких, что для $\forall \omega_k \in \mathrm{B}(m)\ \exists\ d_{km} \in D$. На множестве $\mathrm{A} \cup \Omega$ всех АЭ введем функцию $l(\cdot)$ — номер вышестоящего п-центра по отношению к АЭ, к α_l-агенту и/или ω_k-центру, так что $\alpha_i \in \mathrm{B}(l(i))$ и $\omega_i \in \mathrm{B}(l(i))$. Для ω_u-центра функция $l(\cdot)$ не определена.

В соответствии с общей концепцией комплексной деятельности [4] многоуровневых ОС [2] агенты непосредственно реализуют полезную деятельность, а п-центры организуют агентов и управляют ими. Активность агентов и п-центров заключается в принятии ими в каждом периоде решения об участии в ОС и выборе действия $y_i(t) \in Y_i$ (отказ от участия является частным случаем действия). Действие каждого ω_m-центра заключается в назначении всем нижестоящим АЭ, входящим в m-ю подсистему ($\alpha_i \in \mathrm{B}(m)$ и $\omega_k \in \mathrm{B}(m)$), функций стимулирования и выполнении расчетов согласно функциям стимулирования. Для сохранения общности обозначений действия п-центров будем помечать как $y_i(t) \in Y_i$, понимая под $y_i(\cdot)$ совокупность параметров формируемой системы стимулирования.

Аналогично тому, как это представлено в ДСОС, считаем, что затраты каждого α_i-агента зависят от действий всех агентов в $\omega_{l(i)}$-й подсистеме, к которой принадлежит агент ($\alpha_i \in \mathrm{B}(l(i)) \cap \mathrm{A}$), $c_i^t(y_{\mathrm{B}(i)\,\cap\,\mathrm{A}}[1*t])\colon \prod\limits_{\alpha_k \in \mathrm{B}(m) \cap \mathrm{A}} (Z_k)^t \to \Re_+^1$, а затраты $c_m^t(\cdot)$ каждого ω_m-центра зависят от действий агентов в $\omega_{l(m)}$-й и всех нижестоящих подсистемах $\alpha_k \in \mathrm{A}(l(m))$, $c_m^t(y_{\mathrm{A}(l(m))}[1*t])\colon \prod\limits_{\alpha_k \in \mathrm{A}(l(m))} (Z_k)^t \to \Re_+^1$ (см. рис. 1). Так как $\forall m\ \mathrm{A}(m) \subseteq \mathrm{A}$ и $\{\mathrm{A} \cap \mathrm{B}(m)\} \subseteq \mathrm{A}$, без ограничения общности можно считать, что все функции затрат зависят от действий всех агентов $c_i^t(y_{\mathrm{A}}[1*t])\colon \prod\limits_{\alpha_k \in \mathrm{A}} (Z_k)^t \to \Re_+^1$.

Предположим, что каждая подсистема (и вся система м-ДСОС, рис. 2) в течение каждого периода $t \in \{1, 2, ..., T\}$ функционирует следующим образом.

Шаг 1. В начале периода t каждый ω_m-центр (кроме ω_u-го) последовательно по убыванию технологических номеров получает от вышестоящего для него $\omega_{l(m)}$-го центра информацию о функции стимулирования σ_m^t, после чего принимает решение, участвовать ли в ОС (учитывая функцию $\sigma_m^t(\cdot)$ и функцию своих затрат $c_m^t(\cdot)$). Вышестоящий в м-ДСОС ω_u-центр принимает решение первым на основании своей функции дохода $h_u^t(\cdot)$, заданной на множестве Z_u возможных значений $z_u(t)$ выхода ОС в целом ($h_u^t(z_u[1{*}t])$: $(Z_u)^t \to \Re_+^1$), и своих затрат $c_u^t(\cdot)$.

Шаг 2. Приняв решение участвовать в ОС, ω_m-центр реализует свое действие $y_m(t) \in Y_m$: назначает функции стимулирования всем участникам ω_m-й подсистемы $\sigma_i^t(\cdot)$. Агенты $\alpha_i \in \mathrm{B}(m)$ и п-центры $\omega_k \in \mathrm{B}(m)$ принимают решения об участии в ОС или об отказе от него. Участвующие в ОС агенты выбирают свои действия $y_i(t)$, а ω_k-центры — свои действия $y_k(t)$: назначают системы стимулирования в ω_k-х подсистемах. При отказе ω_m-центра от участия в ОС все $\alpha_i \in \mathrm{A}(m)$ и $\omega_k \in \Omega(m)$ также не участвуют в ОС.

Шаг 3. Последовательно по возрастанию технологических номеров агенты $\alpha_i \in \mathrm{B}(m)$ выбирают и реализуют действия $y_i(t)$, формируются их результаты $z_i(t)$ и результат $z_m(t)$ деятельности каждой из ω_m-подсистем в периоде t. Результат $z_i(t)$ каждого α_i-агента (аналогично ДСОС, [1, § 1]) определяется его действиями $y_i(\cdot)$, результатами его предшественников $\alpha_j \in N(i)$ и $\omega_j \in N(i)$ в сети G_m и его технологической функцией $Q_i^t(\cdot)$ так, что $z_i(t) = Q_i^t(y_i[1{*}t],\ z_{N(i)}[1{*}t])$. Результат ω_m-подсистемы $z_m(t)$ (Z_m — множество его значений, $z_m(t) \in Z_m$) определяется результатами всех элементов ω_m-подсистемы (α_r-х агентов, $\alpha_r \in \mathrm{B}(m)$, и ω_s-х подсистем, $\omega_s \in \mathrm{B}(m)$), результатами «предшественников» в вышестоящей подсистеме $N(l(\omega_m))$ и отображением $z_m(t) = Q_m^t(z_{\mathrm{B}(m)}[1{*}t],\ z_{N(l(m))}[1{*}t])$, назовем его отображением агрегирования $Q_m^t(\cdot)$: $\prod_{\alpha_i \in \mathrm{B}(m) \cup N(l(m))} (Z_i)^t \times \prod_{\omega_k \in \mathrm{B}(m) \cup N(l(m))} (Z_k)^t \to Z_m$.

Шаг 4. В завершение периода последовательно по убыванию технологических номеров каждый ω_m-центр получает компенсацию затрат от вышестоящего, $\omega_{l(\omega_m)}$-го, центра $\sigma_m^t(\cdot)$ (или резервную полезность u_m) и компенсирует затраты всем участникам ω_m-й подсистемы, следуя функциям $\sigma_i^t(\cdot)$.

Шаг 5. Переход к следующему периоду времени $t = t + 1$.

Технологические функции и отображения $Q_m^t(\cdot)$, а также технологические сети G_m и иерархия Γ однозначно определяют зависимость результатов от действий агентов $z_J(y_{\mathrm{A}}[1{*}t])$ для любых множеств J агентов и п-центров: $J \subseteq \mathrm{A} \cup \Omega$.

Будем считать, что каждый ω_m-центр наблюдает фактические действия всех ω_k-центров, входящих в ω_m-ю подсистему, $\omega_k \in \mathrm{B}(m)$, непосредственно нижестоящих по отношению к нему, в текущем периоде, т. е. каждый вышестоящий п-центр знает системы стимулирования, назначаемые каждым из нижестоящих п-центров в пределах их подсистем.

Применительно к агентам предположим два возможных случая — справедливость одного из предположений, П1′ или П2′.

Предположение П1′. Каждый ω_m-центр наблюдает только значение выхода подсистемы $z_m(t)$, но отображения $Q_m^t(\cdot)$ взаимно однозначны относительно результатов α_r-агентов, $\alpha_r \in \mathrm{B}(m)$, и ω_s-х подсистем, $\omega_s \in \mathrm{B}(m)$ (нижестоящих активных элементов), а также α_r-агентов, $\alpha_r \in N(l(m))$, и ω_s-х подсистем, $\omega_s \in N(l(m))$ (его предшественников в вышестоящей подсистеме), в текущем периоде.

Примеры выполнения предположения П1 (см. первую часть статьи [1]) характерны также и для предположения П1′.

Предположение П2′. Каждый ω_m-центр наблюдает фактические действия α_r-агентов, $\alpha_r \in \mathrm{B}(m)$, и ω_s-подсистем, $\omega_s \in \mathrm{B}(m)$ (нижестоящих активных элементов), а также α_r-агентов, $\alpha_r \in N(l(m))$, и ω_s-х подсистем, $\omega_s \in N(l(m))$ (его предшественников в вышестоящей подсистеме), в текущем периоде.

Пусть имеет место информированность участников ОС на момент принятия решений в каждом периоде t (в любом из случаев — справедливости предположения П1 или П2 [1]):

— всем участникам каждой ω_m-й подсистемы (ω_m-центру, α_l-агентам, $\alpha_l \in \mathrm{B}(m)$, и ω_k-центрам, $\omega_k \in \mathrm{B}(m)$), известны множества значений действий Y_l и результатов Z_k, целевые функции $F_i(\cdot)$, включая функции затрат и системы стимулирования, информированность и дальновидность друг друга, а также фактические значения действий

$y_l[1*(t-1)]$ и результатов $z_k[1*(t-1)]$ в предыдущие периоды;

— ω_m-центру известны также отображение $Q_m^t(\cdot)$, множество Σ возможных функций стимулирования и назначенная ему вышестоящим центром функция стимулирования $\sigma_m^t(\cdot)$;

— ω_u-центру известна кроме того функция дохода $h_u(\cdot)$.

При справедливости предположения П1′ ω_m-центр формирует функции стимулирования каждого i-го из нижестоящих агентов на основании $z_m[1*t]$ результата ω_m-й подсистемы $\sigma_i^t(z_m[1*t])$: $(Z_m)^t \to \Re_+^1$, а при выполнении предположения П2′ — на основе непосредственно действий i-го агента так, что $\sigma_i^t(y_i[1*t])$: $(Y_i)^t \to \Re_+^1$. Так как действия ω_m-центров наблюдаемы со стороны вышестоящих п-центров, их функции стимулирования всегда могут быть сформированы на основе их действий $\sigma_m^t(y_m[1*t])$: $(Y_m)^t \to \Re_+^1$.

Тогда:

$$f_i^t(\sigma_i^t, y_A[1*t]) = -c_i^t(y_{A(l(i))}[1*t]) +$$

$$+\begin{cases}\sigma_i^t(z_{l(i)}(y_A[1*t])), \text{если верно П1},\\ \sigma_i^t(y_i[1*t]), \text{если верно П2},\end{cases}$$ — целевая функция α_i-го агента, $\alpha_i \in A$, в t-м периоде;

$$F_i(\{\sigma_i^\tau\}\ \tau = \overline{t, T}, y_A[1*T]) =$$

$$= \sum_{\tau=t}^{T} \delta_i(t, \tau) f_i^\tau(\sigma_i^\tau, y_A[1*\tau])$$ — целевая функция α_i-го агента, $\alpha_i \in A$, $t \in \{1, 2, ..., T\}$;

$$f_m^t(\sigma_m^t, y_A[1*t]) = \sigma_m^t(y_m[1*t]) - c_m^t(y_{A(l(m))}[1*t]) -$$

$$- \sum_{\alpha_i, \omega_i \in B(m)} \sigma_i^t(\cdot)$$ — целевая функция ω_m-го центра, $\omega_m \in \Omega$ ($\omega_m \ne \omega_u$), в t-м периоде;

$$F_m(\{\sigma_m^\tau\}\ \tau = \overline{t, T}, y_A[1*T]) =$$

$$= \sum_{\tau=t}^{T} \delta_m(t, \tau) f_m^\tau(\sigma_m^\tau, y_A[1*\tau])$$ — целевая функция ω_m-го центра, $\omega_m \in \Omega$ ($\omega_m \ne \omega_u$), $t \in \{1, 2, ..., T\}$;

$$\phi_u^t(z_u(y_A[1*t])) = h_u^t(z_u(y_A[1*t])) - c_u^t(y_u[1*t]) -$$

$$- \sum_{\alpha_l, \omega_m \in B(u)} \sigma_i^t(\cdot)$$ — целевая функция ω_u-центра и ОС в целом в t-м периоде;

$$\Phi_u(z_u(y_A[1*t])) = \sum_{t=1}^{T} \delta_u(1, t) \phi_u^t(z_u(y_A[1*t]))$$ — целевая функция ω_u-центра и ОС в целом с учетом распределения дальновидностей $\delta_u(\cdot)$.

Задачу управления сформулируем как построение системы стимулирования (функций $\{\sigma_{A\cup\Omega}^t\}$ для всех п-центров ($\omega_m \in \Omega$) и всех агентов ($\alpha_l \in A$), составляющих ОС), оптимизирующей целевую функцию $\Phi_u(y_{A\cup\Omega}[1*t])$ корневой, ω_u-й, подсистемы:

$$\Phi_u(y_{A\cup\Omega}[1*T]) \to \max_{\{\sigma_{A\cup\Omega}^t\} \in \Sigma}. \quad (11)$$

Аналогично тому, как это было сделано для ДСОС [1, *§ 2*], построим для м-ДСОС компенсаторную систему стимулирования, оптимальную по затратам ω_u-центра и реализующую заданные действия агентов и п-центров.

4. ДЕКОМПОЗИЦИЯ ИГРЫ АГЕНТОВ В МНОГОУРОВНЕВОЙ МНОГОЭЛЕМЕНТНОЙ ДИНАМИЧЕСКОЙ ОРГАНИЗАЦИОННОЙ СИСТЕМЕ

Пусть справедливо предположение П1′. Зафиксируем некоторую допустимую траекторию выхода ОС в целом (ω_u-й подсистемы) $x[1*t]$ — траекторию *планов по результату* системы в целом. Обозначим как $y_A(x[1*t])$ траекторию *планов по действиям* — наборы векторов действий агентов, приводящих к траектории — выходу $x[1*t]$. При справедливости предположения П1′ (аналогично первой части статьи, [1, *§ 2*]) может быть построена наилучшая для центра траектория планов по действиям $y_A(x[1*t])$ для любой траектории $x[1*t]$ планов по результату ОС.

Обозначим $Z_i^t(y) \subseteq Z_{l(i)}$ подмножество множества $Z_{l(i)}$ возможных выходов вышестоящей для α_i-го агента $\omega_{l(i)}$-й подсистемы, $\alpha_i \in B(l(i))$, каждый элемент которого соответствует выходу $z_{l(i)}(t)$ при условии, что i-й агент в t-м периоде выбрал действие y, а остальные агенты и п-центры — любые допустимые для них действия при любых траекториях действий $y_{B(l(i))}[1*(t-1)]$ всех активных элементов в течение предыдущих периодов, т. е.

$$Z_i^t(y) = \{z_k(\{y, y_{-i}(t), y_{B(l(i))}[1*(t-1)]\}):$$
$$\forall y_{-i}(t) \in \prod_{s \in B(l(i)), s \ne i} Y_s,$$
$$\forall y_{B(l(i))}[1*(t-1)] \in \prod_{s \in B(l(i))} (Y_s)^{t-1}\}.$$

Для каждой из возможных траекторий выходов $x[1*t]$ и траекторий $y_{\mathrm{A}}(x[1*t])$ векторов действий всех α_i-х агентов, $\alpha_i \in \mathrm{A}$, построим функцию стимулирования:

$$\hat{\sigma}_i^t(z_{l(i)}[1*t]) = \begin{cases} c_i^t(y_{\mathrm{B}(l(i))}(z_{l(i)}[1*t])) + u_i, \ z_{l(i)}(t) \in Z_i^t(y_i^{\text{п}}(t)), \\ 0, \qquad\qquad\qquad z_{l(i)}(t) \notin Z_i^t(y_i^{\text{п}}(t)). \end{cases} \qquad (12)$$

Для каждой из подсистем $\omega_m \in \Omega$ определим функции $C_m(\cdot)$, выражающие затраты на реализацию траектории выхода подсистемы $z_m[1*t]$, соответствующего $x[1*t]$:

$$C_m^t(y_{\mathrm{A}}(x[1*t])) = c_m^t(y_{\mathrm{A}}(x[1*t])) + u_m + \sum_{\alpha_l \in \mathrm{B}(r)} (c_l^t(y_{\mathrm{A}}(x[1*t])) + u_l) + \sum_{\omega_k \in \mathrm{B}(r)} C_k^t(y_{\mathrm{A}}(x[1*t])), \qquad (13)$$

а на их основе построим функции стимулирования ω_m-центров:

$$\hat{\sigma}_m^t(\{\chi, y_m[1*(t-1)]\}) = \begin{cases} C_m(y_{\mathrm{A}}(x[1*t])), \ \chi = y_m^{\text{п}}(t), \\ 0, \qquad\qquad \chi \neq y_m^{\text{п}}(t), \end{cases} \qquad (14)$$

где χ — фактическое действие ω_m-центра в t-м периоде (система стимулирования, назначенная ω_m-центром всем АЭ в пределах ω_m-й подсистемы).

Покажем, что функции (12)—(14) образуют систему стимулирования, при которой траектории векторов действий агентов и п-центров независимо от их дальновидности составляют иерархию равновесий в каждой из подсистем.

Рассмотрим возможные решения некоторого ω_i-центра в t-м периоде. Пусть $y_{\mathrm{A}\cup\Omega}[1*(t-1)]$ — некоторая история игры, а $y_{-i}(t) = (y_1(t), ..., y_{i-1}(t), y_{i+1}(t), ..., y_{|\mathrm{A}\cup\Omega|}(t))$ — обстановка игры для ω_i-центра (рис. 3).

В начале периода ω_i-центр может получить предложение участвовать в деятельности и получать вознаграждение в соответствии с функцией стимулирования $\hat{\sigma}_i^t(\cdot)$ (9) от вышестоящего $l(\omega_i)$-центра, если все вышестоящие п-центры приняли решение участвовать в ОС, или не получить его (если хотя бы один из вышестоящих п-центров отказался от участия в ОС).

Если ω_i-центр не получает стимулирования, у него нет выбора: он не участвует в ОС и получает резервную полезность u_i.

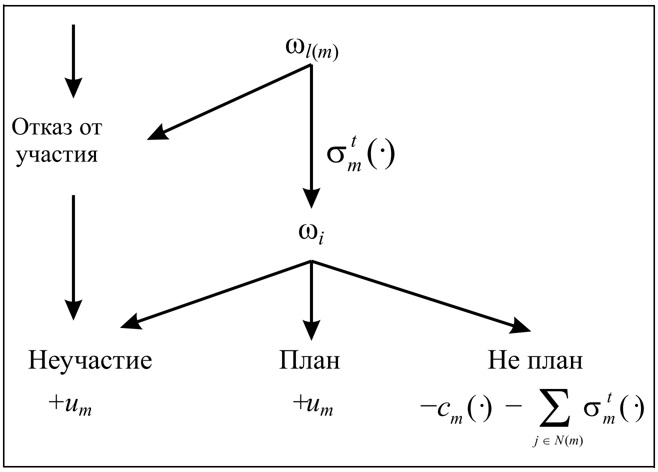

Рис. 3. Возможные решения ω_i-центра в t-м периоде

Если ω_i-центр получает предложение участвовать в ОС на условиях, определяемых соответствующей функцией стимулирования $\sigma_i^t(\cdot)$, у него есть опции:

— назначить систему стимулирования (12)—(14) всем участникам ω_i-й подсистемы, что соответствует его плановому действию $y_i^{\text{п}}(t)$, тогда согласно выражению (14) значение его целевой функции

$$f_i^t(\hat{\sigma}_N^t, y_{\mathrm{A}\cup\Omega}[1*t], t) = \hat{\sigma}_i^t(y_i[1*t]) - c_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) = c_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) + u_i - c_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) = u_i > 0,$$

— отказаться от участия в ОС, тогда значение его целевой функции

$$f_i^t(\hat{\sigma}_N^t, y_{\mathrm{A}\cup\Omega}[1*t]) = u_i > 0,$$

— выбрать действие $s \in Y_i$, отличающееся от планового $s \neq y_i^*(t, x[1*t])$, тогда согласно выражению (12) значение его целевой функции

$$f_i^t(\hat{\sigma}_N^t, y_{\mathrm{A}\cup\Omega}[1*t]) = \hat{\sigma}_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) - c_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) = 0 - c_i^t(y_{\mathrm{A}\cup\Omega}[1*t]) \leq 0.$$

Следовательно, целевая функция п-центра примет вид (совпадающий с формулой (6) из работы [1] с точностью до обозначений):

$$F_i(\{\hat{\sigma}_N^t\}, y_{\mathrm{A}\cup\Omega}[1*T]) = u_i\Delta_i - \sum_{\tau \in \Theta_i} \delta_i(1, \tau)[u_i + c_i^\tau(y_{\mathrm{A}\cup\Omega}[1*\tau])], \quad (15)$$

где Θ_i — множество периодов, в которых ω_i-центр участвовал в ОС, но отклонялся от плановой траектории $y_i^{\text{п}}[1^*T]$ — формирования системы стимулирования (12)—(14).

Для каждого из α_i-агентов, повторив выкладки из работы [1, *§ 2*], можно показать, что целевая функция агента примет вид, с точностью до обозначений совпадающий с формулами (6) и (15):

$$F_i(\{\hat{\sigma}_N^t\},\ y_{\text{A}}[1^*T]) =$$
$$= u_i\Delta_i - \sum_{\tau \in \Theta_i} \delta_i(1,\ \tau)[u_i + c_i^{\tau}(y_{\text{A}}[1^*\tau])],$$

где Θ_i — множество периодов, в которых α_i-агент участвовал в ОС, но отклонялся от плановой траектории $y_i^{\text{п}}[1^*T]$.

Таким образом, при соблюдении гипотезы благожелательности оптимальная стратегия каждого активного элемента ОС — ω_i-центра и/или α_l-агента — заключается в выборе планового для него действия в каждом периоде, когда он получает контракт от вышестоящего п-центра, независимо от любых иных действий остальных участников. При неполучении контракта он не участвует в игре, а реализует резервную полезность.

По аналогии с доказательством утверждения 1 из работы [1] несложно показать, что никакие затраты, меньшие $C_u^t(y_{\text{A}}(x[1^*t]))$ (см. выражение (13)), не обеспечат реализации траектории $y_{\text{A}}(x[1^*t])$, поэтому справедливо

Утверждение 3. *Если выполнено предположение П1', то система стимулирования* (12)—(14) *реализует траекторию планов по действию* $y_N(x[1^*t])$ *как РДС игры агентов с минимальными затратами центра на стимулирование. Выигрыши всех агентов в каждом периоде в этом равновесии тождественно равны их резервным полезностям. Затраты центра при этом*

$$C(x[1^*T]) = \sum_{t=1}^{T} \delta(1,\ t) \times$$
$$\times \sum_{\alpha_i \in \text{A};\ \omega_i \in \Omega} (c_i^t(y_{\text{A}\cup\Omega}(x[1^*t])) + u_l). \ \blacklozenge \qquad (16)$$

В случае справедливости предположения П2′ функции стимулирования (12) примут вид (17):

$$\hat{\sigma}_i^t(y_{\text{A}}^{\text{п}}[1^*t]) =$$
$$= \begin{cases} c_i^t(y_{\text{B}(l(i))}[1^*t]) + u_i, & y_i(t) = y_i^{\text{п}}(t), \\ 0, & y_i(t) \neq y_i^{\text{п}}(t), \end{cases} \qquad (17)$$

и доказательство утверждения 3 может быть повторено с точностью до обозначений, откуда следует

Утверждение 4. *Если выполнено предположение П2', то система стимулирования* (13), (14) *и* (17) *реализует траекторию планов по действию* $y_N(x[1^*t])$ *как РДС игры агентов с минимальными затратами центра на стимулирование. Выигрыши всех агентов в каждом периоде в этом равновесии тождественно равны их резервным полезностям. Затраты центра при этом вычисляются по формуле* (16). ♦

В случае невыполнения гипотезы доброжелательности аналогично тому, как это было сделано для ДСОС в работе [1, *§ 2*], выражения (12)—(17) должны быть дополнены стимулирующими надбавками по аналогии с выражениями (9) и (10) из работы [1].

Итак, в многоуровневой многоэлементной динамической ОС с ограничением на совместную деятельность агентов в каждой подсистеме в виде технологических сетей для любой осуществимой траектории выхода ОС $z_u[1^*T]$ может быть построена компенсаторная система стимулирования, которая:

— реализует как РДС траекторию действий агентов и п-центров,

— декомпозирует задачу по агентам, п-центрам и по периодам,

— обеспечивает гарантированно (по всем возможным дальновидностям агентов и п-центров) минимальные затраты центра $C_u^t(y_{\text{A}}(x[1^*t]))$ на реализацию этой траектории, определяемые формулой (16).

Как и для ДСОС, отметим, что при соблюдении предположений П1′ или П2′ и других условий задачи центру удается построить такую оптимальную систему стимулирования, которая обеспечивает равновесие в доминантных стратегиях игры агентов независимо от их взаимосвязей — технологии м-ДСОС (технологических функций и отображений $Q_m^t(\cdot)$, а также технологических сетей G_m и иерархии Г).

В Приложении показано, что полученные результаты в детерминированной постановке распространяются на модели ОС, в которых присутствует интервальная или вероятностная неопределенность относительно затрат агентов.

5. ПЛАНИРОВАНИЕ В ДИНАМИЧЕСКИХ СЕТЕВЫХ АКТИВНЫХ СИСТЕМАХ

Общая схема решения задач управления ОС с *побочными платежами* (стимулирование игроком, делающим первый ход, игроков, делающих ход вторыми, в теории иерархических игр в общем случае называется игрой Γ_2 с побочными платежами [5]) такова [6]: сначала для каждого допустимого вектора действий агентов (или вектора результатов их деятельности) согласно принципам декомпозиции и компенсации затрат агентов решается задача *согласования* — поиска компенсаторной системы стимулирования, реализующей эти действия/результат с минимальными затратами центра. Затем на втором шаге решается задача *оптимального планирования* — поиска вектора действий/результатов, реализация которого наиболее выгодна для центра. Отметим, что такая двушаговая схема значительно упрощает решение задач (3) и (11) (в которых необходимо искать функции, позволяющие функционалу, вычисляемому на множестве, зависящем от искомой функции, достичь экстремума), сводя ее к нахождению относительно простой системы стимулирования и решению затем задачи «скалярной» оптимизации.

Доказанные в первой [1] и настоящей частях статей утверждения 1—4 обеспечивают применимость принципа декомпозиции игры агентов в динамических и многоуровневых ОС (в том числе с затратами агентов, характеризующимися неопределенностью). Для решения задач согласованного управления в таких организационных системах остается найти оптимальные планы. Поэтому рассмотрим теперь решение задач оптимального планирования для ДСОС и м-ДСОС.

Построенные в § 2 и 4 компенсаторные системы стимулирования (4), (7) и (12)—(14), (17) позволяют декомпозировать задачи (3) и (11) соответственно по периодам и участникам и свести их к задачам согласованного планирования, что полностью соответствует общей концепции решения задач управления, принятой в теории активных систем [7, 8] и теории управления организационными системами [6]. Укажем алгоритм решения задач (3) и (11).

При справедливости предположения П1 или П1′ субъект (центр), осуществляющий управление в ДСОС или м-ДСОС, может решить задачу планирования — найти оптимальный план $x^*[1^*T]$ (траекторию выхода сети), максимизирующий разность между выручкой и суммарными затратами всех агентов с учетом распределения дальновидностей центра:

— для ДСОС:

$$\sum_{t=1}^{T}\delta(1,\ t)\{h(x[1^*t]) - \sum_{i=1}^{n}(c_i(y_N^*(x[1^*t])) + u_i)\} \to \max_{x[1^*T]}, \qquad (18)$$

— для м-ДСОС:

$$\sum_{t=1}^{T}\delta(1,\ t)\{h_u(x[1^*T]) - C_u(z_u(x[1^*T]))\} \to \max_{x[1^*T]}. \qquad (19)$$

При выполнении предположения П2 или П2′ задачи планирования имеют вид:

— для ДСОС:

$$\sum_{t=1}^{T}\delta(1,\ t)\{h(z_n(y_N[1^*t])) - \sum_{i=1}^{n}(c_i(y_N[1^*t]) + u_i)\} \to \max_{y_N[1^*T]}, \qquad (20)$$

— для м-ДСОС:

$$\sum_{t=1}^{T}\delta(1,\ t)\{h_u(z_u(y_N[1^*t])) - C_u(z_u(y_N[1^*t]))\} \to \max_{y_N[1^*T]}. \qquad (21)$$

Задачи типа (18)—(21) нетривиальны (могут быть решены численно, а в частных случаях — методом динамического программирования), но для их решения необходимо искать вектор планов, а не вектор-функцию, как в задачах (3) или (11).

Если агенты полностью дальновидны, и центр знает распределения их дальновидностей[4] $\delta_i(1,\ t)$, он может воспользоваться этим и повысить эффективность своего управления $\Phi(\sigma_N)$, применяя *аккордную систему стимулирования*. Для ДСОС в случае справедливости предположения П1 аккордная система определяется как

$$\sigma_i^t(z_n[1^*t]) = \begin{cases} \tilde{C}_i(z_n[1^*T]), \text{ при } z_n(\tau) \in Z_i^t(y_i^{\text{п}}(\tau)) \\ \qquad \text{для всех } \tau \in [1,\ T] \text{ и } t = T, \\ 0, \text{ при } z_n(\tau) \notin Z_i^t(y_i^{\text{п}}(\tau)) \text{ хотя бы для} \\ \qquad \text{одного } \tau \in [1,\ t] \text{ или } t < T, \end{cases}$$

где $\tilde{C}_i(z_n[1^*t]) = \delta_i^{-1}(1,\ T)\sum_{t=1}^{T}\delta_i(1,\ t)(c_i^t(y_N(z_n[1^*t])) + u_i)$, а траектория оптимального плана получена из условия $\left\{\sum_{t=1}^{T}\delta(1,\ t)h(x[1^*t]) - \delta(1,\ T)\sum_{i=1}^{n}\tilde{C}_i(x[1^*t])\right\} \to \max_{x[1^*T]}$ и может отличаться от траектории, полу-

---

[4] И вся эта информация является общим знанием всех игроков.

чаемой из выражения (18) в силу возможных различий распределений дальновидностей центра и агентов. Для м-ДСОС аккордная система может быть получена аналогично.

Дополнительный выигрыш от применения аккордной системы возникает у центра не всегда и зависит от соотношения распределений дальновидностей центра и агентов. В частности, в случае совпадающих распределений дальновидностей (например, $\delta_i(1,\ t) = \delta(1,\ t) = \exp(-\alpha(t-1))$, где $\alpha \geq 0$), эффективность управления для обоих случаев совпадает.

Таким образом, задачи оптимального планирования (18) и (21) дополняют решение задач стимулирования — компенсаторные системы стимулирования (4), (7) и (12)—(14), (17) — и обеспечивают решение задач согласованного управления многоэлементными динамическими организационными системами (3) и (11).

ЗАКЛЮЧЕНИЕ

Обсудим полученные результаты. Отметим, что компенсаторные системы стимулирования, декомпозирующие задачи по агентам и периодам, удалось построить для случая достаточно сложных моделей организационных систем — одновременно многоэлементных и динамических.

Основная особенность постановок задачи, позволившая это сделать, заключается в предположении П1 о взаимной однозначности технологических функций относительно действий агентов и результатов их предшественников в текущем периоде (или в предположении П2 о полной наблюдаемости центром действий агентов). Это позволяет не только последовательно по периодам строить желательные траектории действий агентов в зависимости от траектории плана по «выходу всей ОС», но также контролировать действия агентов, наблюдая только лишь фактические значения «выхода всей ОС». Поэтому в условиях таких жестких технологических зависимостей, условно говоря, стимулирование за результат эквивалентно стимулированию за действие. Также заметим, что все приведенные выкладки и полученные результаты допускают расширение на случай, когда множества возможных значений действий агентов и их результатов Y_i и Z_i изменяются во времени и зависят от предыдущих действий всех агентов известным образом.

Интересная особенность полученных решений состоит в том, что жесткие технологические зависимости позволили декомпозировать задачу не только по агентам, но и по периодам. Возможность такой декомпозиции обычно характерна для прямо противоположного случая — полного отсутствия связей: независимых агентов и несвязанных периодов. Здесь же возможность декомпозиции обеспечена именно благодаря взаимной однозначности обобщенной технологии деятельности агентов.

Отметим также существенную практическую значимость полученных результатов. Рассмотренная постановка согласованного управления иерархической совокупностью организационных систем, технологически связанных между собой, естественным образом отражает проблему управления бизнесом современных крупных предприятий, который требует организации кооперации и управления десятками тысяч относительно автономных бизнес-элементов, часть из которых, в свою очередь, являются предприятиями. Важно, что получены результаты как в детерминированной постановке, так и для моделей ОС, в которых затраты участников характеризуются неопределенностью, допускающей интервальное или вероятностное представление.

Доказанные в работе утверждения разделяют математические модели и методы управления сложными (многоэлементными, многоуровневыми динамическими) организационными системами, с одной стороны, на теоретико-игровые задачи, учитывающие активность участников ОС, и, с другой стороны, — на оптимизационные задачи планирования (в том числе распределения ресурсов). Независимость систем стимулирования и от технологических графов, и от технологических функций позволяет рассматривать отдельно задачи управления технологией деятельности. Все это позволяет корректно структурировать единую проблему управления на три группы математических моделей и задач (поиск технологии, планирование и стимулирование) и решать их по отдельности, учитывая взаимосвязи.

ПРИЛОЖЕНИЕ

Модели ОС, в которых присутствует неопределенность относительно затрат агентов

Рассмотрим модели организационных систем, в которых затраты агентов характеризуются интервальной или вероятностной неопределенностью. Данные модели имеют большое прикладное значение, так как на практике реализация любой, даже многократно повторяющейся деятельности связана с возможностью наступления априори неопределенных событий, порождаемых как факторами внешней среды, так и особенностями субъекта и/или предмета в конкретные моменты времени [4]. При этом деятельность, как правило, выполняется до достижения априори поставленной цели, а наступление событий неопределенности вызывает неопределенность затрат субъекта (по достижению цели). Распространим приведенные в первой [1] и второй частях статьи результаты на ОС с неопределенными затратами агентов, причем будем рассматривать одновременно постановки и ДСОС (§ 1 и 2), и м-ДСОС (§ 3 и 4).

Будем считать имеющими место все обозначения и определения величин, множеств, функций и функционалов t, $z_i(\cdot)$, $y_i(\cdot)$, u_i, Y_i, Z_i, D, $N(\cdot)$, $E(\cdot)$, $\mathrm{A}(\cdot)$, $\Omega(\cdot)$, $\mathrm{B}(\cdot)$, $Q_i^t(\cdot)$, $\sigma_i^t(\cdot)$, $\delta(\cdot)$, $f_i^t(\cdot)$, $F_i(\cdot)$, $\phi_u^t(\cdot)$, $C(\cdot)$, $\Phi(\cdot)$, введенные в § 1—4, кроме функции затрат $c_i^t(\cdot)$. Размер затрат, требуемых от каждого i-го агента и/или п-центра для выбора и реализации действия в любом периоде t, будем считать неопределенным на момент принятия участниками ОС решений и обозначать через $\tilde{c}_i(t)$. Предполагаем, что апостериори затраты становятся известными агентам, но остаются неизвестными центрам (если апостериори затраты известны центрам, то задача сводится к детерминированному случаю, рассмотренному в § 1—4).

Интервальная модель неопределенности. Пусть для затрат справедливы условия:

$$\forall\ t,\ i\colon\ c_i^t(y_N[1{*}t]) \ge \tilde{c}_i(t), \tag{22}$$

т. е. будем считать, что фигурирующие в моделях (см. § 1—4) функции $c_i^t(\cdot)\colon\ \prod_{k\in N}(Y_k)^t \to \Re_+^1$ задают максимальные значения возможных затрат i-го агента и/или п-центра по выбору и реализации действия в любом периоде t.

Если функции $c_i^t(\cdot)$ и неравенства (22) являются общим знанием[5] всех участников ОС на момент выбора ими действий в каждом периоде, то неопределенность затрат может быть устранена с помощью метода гарантированного результата.

Задача управления в этом случае будет заключаться в формировании системы стимулирования $\{\sigma_N^t\}$, имеющей максимальную гарантированную эффективность — обеспечивающую максимум целевой функцией центра $\Phi(\cdot)$ при наихудших значениях затрат $\zeta_i(\cdot)$ всех агентов и п-центров во всех периодах:

$$\min_{\{\varsigma_i(\cdot)\}}\{\Phi(\{\sigma_N^t\})\} \to \max_{\sigma_N^t} \quad \text{для ДСОС},$$

$$\min_{\{\varsigma_i(\cdot)\}}\{\Phi_u(y_{\mathrm{A}\cup\Omega}[1{*}T])\} \to \max_{\{\sigma_{\mathrm{A}\cup\Omega}^t\}\in\Sigma} \quad \text{для м-ДСОС}.$$

Логические построения и выкладки, изложенные выше, могут быть повторены с точностью до формулировок и семантики функций $c_i^t(\cdot)$, и таким образом можно сформулировать ряд утверждений, идентичных утверждениям 1—4 и определяющих системы стимулирования $\{\sigma_N^t\}$, совпадающие с выражениями (4), (7), (12)—(14) и (17) с учетом того, что функции $c_i^t(\cdot)$ имеют смысл максимальных значений затрат. Данные системы стимулирования реализуют траектории планов по действию как РДС игры агентов с минимальными затратами центра на стимулирование. Выигрыши всех агентов в каждом периоде в этих равновесиях будут не меньше их резервных полезностей, а затраты центра будут определяться соответствующими выражениями (5), (8), (16).

Вероятностная модель неопределенности. Пусть затраты $\tilde{c}_i(t)$ для всех t и i допускают вероятностное описание — представление их как случайных величин с некоторыми функциями распределения $F_c(\cdot;\, y_N[1{*}t])$, зависящими от $y_N[1{*}t]$ как от параметров, такими, что

$$E[\tilde{c}_i(t)] = \int_z z\,dF_c(z,\, y_N[1{*}\mathrm{t}]) = c_i^t(y_N[1{*}t]), \tag{23}$$

т. е. фигурирующие в моделях (см. § 1—4) функции $c_i^t(\cdot)$: $\prod_{k\in N}(Y_k)^t \to \Re_+^1$ определяют математические ожидания затрат i-го агента и/или п-центра по выбору и реализации действия в любом периоде t.

Если функции $c_i^t(\cdot)$ (знание самих функций распределения $F_c(\cdot;\, y_N[1{*}t])$ не является необходимым) и условия (23) являются общим знанием всех участников ОС на момент выбора ими действий в каждом периоде, то неопределенность относительно затрат может быть устранена с помощью принципа ожидаемой полезности (см. также сноску[4]).

Предполагаем, что агент выбирает свое действие, не зная его априори и исходя как бы из того, что реализуется некоторое ожидаемое значение функции затрат. При этом в отдельных случаях (в зависимости от реализовавшегося значения случайного параметра) поощрение может не компенсировать реализовавшиеся затраты для выбранного агентом действия, а в отдельных случаях — превышать эти затраты. Агент в этом случае считает, что на достаточно длинном интервале времени эти расхождения компенсируются, и в среднем реализуется математическое ожидание функции затрат по случайному параметру.

Задача управления в этом случае будет поставлена как формирование системы стимулирования $\{\sigma_N^t\}$, имеющей максимальную эффективность — обеспечивающей максимум математического ожидания целевой функцией центра:

$$E[\Phi(\{\sigma_N^t\})] \to \max_{\sigma_N^t} \quad \text{для ДСОС},$$

$$E[\Phi_u(y_{\mathrm{A}\cup\Omega}[1{*}T])] \to \max_{\{\sigma_{\mathrm{A}\cup\Omega}^t\}\in\Sigma} \quad \text{для м-ДСОС}.$$

Логические построения и выкладки, изложенные в § 1—4, могут быть повторены с точностью до формулировок и семантики функций $c_i^t(\cdot)$, т. е. можно сформулировать ряд утверждений, идентичных утверждени-

[5] Модели ОС с асимметричной информированностью участников, рассматриваемые теорией контрактов (например, [9—11]) и теорией управления организационными системами [6], безусловно, представляют большой интерес, однако в данной работе не рассматриваются, являясь предметом дальнейших исследований.

ям 1—4 и определяющих системы стимулирования $\{\sigma_N^t\}$, совпадающие с утверждениями (4), (7), (12)—(14) и (17) с учетом, что функции $c_i^t(\cdot)$ имеют смысл математических ожиданий затрат. Такие системы стимулирования реализуют траектории планов по действию как РДС игры агентов с минимальными затратами центра на стимулирование. Математические ожидания выигрышей всех агентов в каждом периоде в этом равновесии будут равны их резервным полезностям, а затраты центра будут определяться соответствующими выражениями (5), (8), (16).

ЛИТЕРАТУРА

1. *Белов М.В.* Согласованное управление многоэлементными динамическими организационными системами. Ч. 1. Динамическая организационная система в составе одного центра и множества агентов // Проблемы управления. — 2020. — № 1. — С. 39—47. [*Belov, M.V.* Incentive-compatible Control in Dynamic multi-agent systems. Part 1. Contracts in dynamic system with one principal and multiple agents // Control Sciences. — 2020. — No. 1. — P. 38—47. (In Russian)]
2. *Новиков Д.А.* Механизмы функционирования многоуровневых организационных систем. — М.: Фонд «Проблемы управления», 1999. — 161 с. [*Novikov, D.A.* Mekhanizmy funktsionirovaniya mnogourovnevykh organizatsionnykh sistem. — Moscow: Fond «Problemy upravleniya», 1999. — 161 s. (In Russian)]
3. *Бурков В.Н., Горгидзе И.А., Ловецкий С.Е.* Прикладные задачи теории графов. — Тбилиси: ВЦ АН ГССР, 1974. — 232 с. [*Burkov, V.N., Gorgidze, I.A., Lovetskii, S.E.* Prikladnye zadachi teorii grafov. — Tbilisi: VTs AN GSSR, 1974. — 232 s. (In Russian)]
4. *Белов М.В., Новиков Д.А.* Методология комплексной деятельности. — М.: Ленанд, 2017. — 310 с. [*Belov, M.V., Novikov, D.A.* Metodologiya kompleksnoi deyatel'nosti. — Moscow: Lenand, 2017. — 310 s. (In Russian)]
5. *Гермейер Ю.Б.* Игры с непротивоположными интересами. — М.: Наука, 1976. — 327 с. [*Germeier, Yu.B.* Igry s neprotivopolozhnymi interesami. — Moscow: Nauka, 1976. — 327 s. (In Russian)]
6. *Новиков Д.А.* Теория управления организационными системами: 3-е изд. — М.: Физматлит, 2012. — 604 с. [*Novikov, D.A.* Teoriya upravleniya organizatsionnymi sistemami: 3-e izd. — Moscow: Fizmatlit, 2012. — 604 s. (In Russian)]
7. *Ашимов А.А., Бурков В.Н., Джапаров Б.А., Кондратьев В.В.* Согласованное управление активными производственными системами. — М.: Наука, 1986. — 248 с. [*Ashimov, A.A., Burkov, V.N., Dzhaparov, B.A., Kondrat'ev, V.V.* Soglasovannoe upravlenie aktivnymi proizvodstvennymi sistemami. — Moscow: Nauka, 1986. — 248 s. (In Russian)]
8. *Бурков В.Н., Кондратьев В.В.* Механизмы функционирования организационных систем. — М.: Наука, 1981. — 384 с. [*Burkov, V.N., Kondrat'ev, V.V.* Mekhanizmy funktsionirovaniya organizatsionnykh sistem. — Moscow: Nauka, 1981. — 384 s. (In Russian)]
9. *Подколозина Е.А., Рябинина А.Ю., Юдкевич М.М.* Основы теории контрактов: Модели и задачи. — М.: ГУ ВШЭ, 2002. — 352 с. [*Podkolozina, E.A., Ryabinina, A.Yu., Yudkevich, M.M.* Osnovy teorii kontraktov: Modeli i zadachi. — Moscow: GU VShE, 2002. — 352 s. (In Russian)]
10. *Bolton, P., Dewatripont, M.* Contract Theory. — Cambridge: MIT Press, 2005. — 740 p.
11. *Salanie, B.* The Economics of Contracts. — Cambridge: MIT Press, 2005. — 224 p.

Статья представлена к публикации членом редколлегии В.Н. Бурковым.

Поступила в редакцию 11.02.2019, после доработки 26.09.2019.
Принята к публикации 24.10.2019.

Белов Михаил Валентинович — д-р техн. наук, компания ИБС, г. Москва, ✉ mbelov59@mail.ru.

INCENTIVE-COMPATIBLE CONTROL IN DYNAMIC MULTI-AGENT SYSTEMS.
Part 2. Contracts in dynamic hierarchical multi-agent system

M.V. Belov

IBS company, Moscow, Russia
✉ mbelov59@mail.ru

Abstract: The results obtained in the Part 1 of the paper are extended to a multi-level dynamic multi-agent system, as well as to the case of uncertain agent costs. It has been proved that for any admissible trajectory of results, a coordinated compensatory incentive system can be constructed that implements (as an equilibrium in dominant strategies) the trajectory of the agents leading to the desired trajectory of results; decomposes the control task by agents and by time periods; provides guaranteed (for all possible far-sighted agents) minimum costs of the governing body of the principal for the implementation of this trajectory of results. It is shown that in such incentive systems, the values of payments depend only on the corresponding values of the cost functions, which, in turn, indirectly take into account the technological functions, network structure and active system structure as a whole. The problem of optimal planning is posed and an algorithm for solving it is indicated.

Keywords: contract theory, incentive problem, dynamic hierarchical multi-agent system.