

НЕЧЕТКАЯ КОНЕЧНО-АВТОМАТНАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОГО МУЛЬТИМОДАЛЬНОГО ИНТЕРФЕЙСА¹

В.В. Девятков, А.Н. Алфимцев

Предложен метод организации мультимодального интеллектуального интерфейса пользователя со средой в процессе распознавания возникающих в ней динамических ситуаций, на основе совокупности взаимодействующих нечетких конечных автоматов. Автоматы служат для задания отношений между модальными объектами, между мультимедийными объектами и между модальными и мультимедийными объектами.

Ключевые слова: модальность, интеллектуальный мультимодальный интерфейс, модальные модели, мультимедийные модели, нечеткий конечный автомат.

ВВЕДЕНИЕ

Развитие охранных, географических, игровых, бытовых, медицинских, транспортных и других информационных систем требует создания таких интеллектуальных интерфейсов, которые позволили бы интуитивно и естественно управлять этими системами, а также дали бы возможность анализировать сложные пространственно-временные ситуации, возникающие в области их применения [1–3]. Но остаются открытыми вопросы о том, какими должны быть формальные модели интеллектуального интерфейса, способные учитывать пространственно-временные отношения модальных объектов и мультимедийных объектов среды, и как их построить, затратив минимальные усилия на проектирование и обучение.

Данная статья следует парадигме Р.А. Брукса [4] о представлении модальных и мультимедийных объектов иерархией взаимосвязанных конечных автоматов, учитывающих отношения между объектами для анализа их поведения напрямую, подавая на них реальные или виртуальные воздействия и получая в ответ реакцию, по которой можно судить об их поведении.

Основная идея настоящей работы состоит в том, что, начиная с некоторого начального (нулевого) уровня, выделяется множество модальных и мультимедийных объектов (далее будем говорить просто объектов) и их атомарных признаков. Признаки используются для синтеза автоматов нулевого уровня. Функция выходов каждого автомата нулевого уровня определяет значение только одного признака уровня в какой-либо момент времени. На основе автоматов нулевого уровня создаются автоматы первого уровня, задающие отношения между признаками нулевого уровня. Каждый автомат первого уровня задает отношения между одним и тем же множеством признаков нулевого уровня. Функция выходов каждого автомата нулевого уровня определяет значение только одного признака нулевого уровня в один момент времени. Каждый автомат последующего уровня u задает отношения между одним и тем же множеством признаков уровней $u' < u$. Функция выходов каждого автомата уровня u определяет значение только одного отношения первичных признаков уровня $u' < u$ в один момент времени. Распознавание динамических ситуаций сводится к установлению сходства входных динамических пространственно-временных отношений объектов с эталонными.

Ориентация статьи на качественные отношения объясняется тем, что они позволяют:

- в естественной для человека манере воспринимать мир на интуитивном уровне [5], это осо-

¹ Работа выполнена при финансовой поддержке гранта Президента РФ МК2010 (проект № МК-3013.2010.9).

бенно полезно при интеллектуальном интерфейсе «человек—компьютер»;

- рассуждать в условиях, когда количественные признаки виртуальных объектов не требуются или невозможны [6].

Качественным пространственным представлениям в последнее десятилетие уделялось много внимания, известны соответствующие обзоры [7]. Было развито много пространственных исчислений, использующих топологические [8] или позиционные [9] рассуждения, а также рассуждения на интервалах [6]. Известны программные реализации методов пространственных рассуждений [10]. Работ, учитывающих одновременно как пространственные, так и временные изменения, сравнительно немного [11—13], их прикладное значение невелико, в них используется только четкое представление отношений.

1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Модальность — присущая человеку форма воздействия на другого человека или компьютер с помощью речи, жестов, прикосновений, мимики, текстов и др.

Модальные объекты — физические объекты (например, кисти рук) или сигналы (например, звуковые), служащие источниками модальностей человека. *Модальные модели* — формальные модели, используемые для представления модальных объектов.

Мультимодальный интерфейс — интерфейс между человеком и компьютером, использующий несколько модальностей.

Мультимедийные объекты — объекты среды, определяемые множеством параметров, зависящих от времени и пространства. В совокупности все эти объекты будем называть *мультимедийной информацией*.

Мультимедийные модели — формальные модели, используемые для представления мультимедийных объектов.

Интеллектуальность мультимодального интерфейса связана со способностью человека использовать модальности в процессе анализа различных ситуаций, возникающих в среде мультимедийных объектов, а также с наличием у него ряда функций, например, общения, обоснования, обучения [14]. Архитектура интеллектуального мультимодального интерфейса (ИМИ) зависит от модальных и мультимедийных моделей и принципов распознавания ситуаций на этих моделях. Взаимодействие этих моделей с ИМИ показано на рис. 1.

Введем несколько основных обозначений. Дискретную шкалу виртуальных моментов времени обоз-

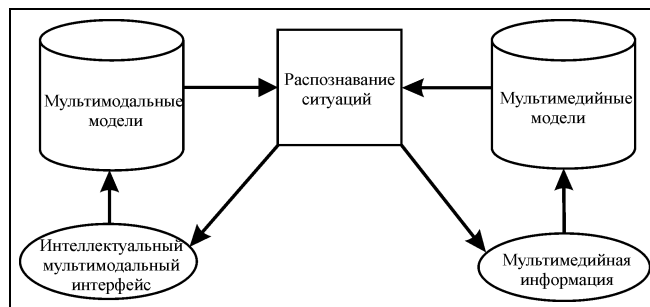


Рис. 1. Место интеллектуального мультимодального интерфейса для анализа мультимедийной информации

начим как $T = \{t | t \in N\}$, где разность $\Delta t = t_{i+1} - t_i$ между двумя соседними отсчетами времени постоянная. Определим на шкале T временной интервал $[t_s, t_e] = \{t | t_s \leq t \leq t_e\}$. Будем полагать, что каждый признак нулевого уровня какого-либо объекта θ в момент времени t может принимать значения $y_{i_1}(\theta_t)$, $i_1 = 1, \dots, m_1$, которые назовем *отсчетами*. Кортеж отсчетов $Y_{i_1}[\theta_{t_s}, \theta_{t_e}] = \langle y_{i_1}(\theta_{t_s}), \dots, y_{i_1}(\theta_{t_e}) \rangle$, $i_1 \in \{1, \dots, m_1\}$ одного признака нулевого уровня в течение нескольких подряд идущих моментов времени $t_s, \dots, t_e$ (в течение временного интервала $[t_s, t_e]$) назовем *трендом*. Множество отсчетов $K(\theta_t) = \{y_1(\theta_t), \dots, y_{m_1}(\theta_t)\}$, принадлежащих m_1 различным трендам в момент времени t назовем *срезом*. Кортеж срезов $\mathcal{K}(\theta) = \langle K(\theta_{t_s}), \dots, K(\theta_{t_e}) \rangle$ назовем *поток срезом*. Совокупность трендов $\mathcal{U}(\theta) = \{Y_1[\theta_{t_s}, \theta_{t_e}], Y_2[\theta_{t_s}, \theta_{t_e}], \dots, Y_n[\theta_{t_s}, \theta_{t_e}]\}$ назовем *поток трендов*.

Будем полагать, что каждое отношение Θ уровня $u \geq 1$ объектов множества $\{\theta^1, \theta^2, \dots, \theta^l\}$, задаваемое на множестве признаков уровня $v < u$, может принимать множество значений $y_{i_u}(\Theta_t)$, $i_u = 1, \dots, m_u$, называемых признаками уровня u или *макроотсчетами*. *Макротрендом* будем называть кортеж $Y_{i_u}[\Theta_{t_s}, \Theta_{t_e}] = \langle y_{i_u}(\Theta_{t_s}), \dots, y_{i_u}(\Theta_{t_e}) \rangle$, $i_u \in \{1, \dots, m_u\}$. *Макросрезом* будем называть множество $K(\Theta_t) = \{y_1(\Theta_t), \dots, y_{m_u}(\Theta_t)\}$. *Макропоток макросрезом* будем называть кортеж $\mathcal{K}(\Theta) = \langle K(\Theta_{t_s}), \dots, K(\Theta_{t_e}) \rangle$. *Макропоток макротрендов* будем называть множество $\mathcal{U}(\Theta) = \{Y_1[\Theta_{t_s}, \Theta_{t_e}], Y_2[\Theta_{t_s}, \Theta_{t_e}], \dots, Y_n[\Theta_{t_s}, \Theta_{t_e}]\}$.

Далее будут рассмотрены пространственные качественные отношения объектов на срезах (стати-



ческие отношения) и пространственные качественные отношения объектов на трендах (динамические отношения), а также отношения на потоках трендов и потоках срезов, представляемые четкими и нечеткими конечными автоматами.

2. ПРИНЦИПЫ ЗАДАНИЯ ОТНОШЕНИЙ

2.1. Отношения на макросрезах

Рассмотрим отношения на макросрезах в двумерном декартовом пространстве, используя понятие кадра $I_t(V, W)$, хранящегося в видеофайле или формируемого цифровой видеокамерой в момент времени t и имеющий число пикселей V и W соответственно по горизонтали и вертикали. Кадр может отображать как модальные, так и мультимедийные объекты. Мультимодальный интерфейс имеет дело не с физическими объектами, а с объектами, являющимися их отображениями и содержащимися в кадре. Общее число пикселей в кадре $I_t(V, W)$ равно $V \times W$. В качестве начала координат примем левый нижний угол кадра. Пронумеруем пиксели по вертикали и горизонтали соответственно от 1 до V и от 1 до W . Множество пикселей кадра $I_t(V, W)$, очерчивающих интересующий нас объект θ , обычно называется *областью интересов*. Область интересов может быть прямоугольной, эллиптической, круглой или какой-либо иной. Обозначим $O_t(X, Y)$ прямоугольную область интересов, захватываемую в кадре $I_t(V, W)$ в момент времени t и содержащую множество пикселей X по вертикали и Y по горизонтали. Общее число пикселей в области интересов $O_t(X, Y)$ равно $|X| \times |Y|$. Отдельный пиксель в кадре обозначим $I_t(v, w)$, $v \in \{1, \dots, V\}$, $w \in \{1, \dots, W\}$, а в области интересов — $O_t(x, y)$. Признаками объекта θ (точнее, отображения объекта, каковым является область интересов) могут быть, например, геометрические параметры этого объекта, координаты центра тяжести, площадь области интересов, длина ее диагоналей и т. п.

Для представления сути предлагаемого подхода ограничимся отношениями $\Theta = \{\theta_t^1, \theta_t^2\}$ на множестве двух объектов θ_t^1, θ_t^2 . Каждый объект имеет соответственно признаки нулевого уровня $y_{11}(\theta_t^1) = x_{0(t)}^1$, $y_{12}(\theta_t^1) = y_{0(t)}^1$, $y_{13}(\theta_t^1) = \alpha_{0(t)}^1$, $y_{21}(\theta_t^2) = x_{0(t)}^2$, $y_{22}(\theta_t^2) = y_{0(t)}^2$, $y_{23}(\theta_t^2) = \alpha_{0(t)}^2$, обозначающие соответственно координаты центров тяжести объектов в глобальной системе координат x_0, y_0 с началом, находящимся в левом нижнем углу кадра, и углы поворота по часовой стрелке осей локальных систем координат объектов вокруг цен-

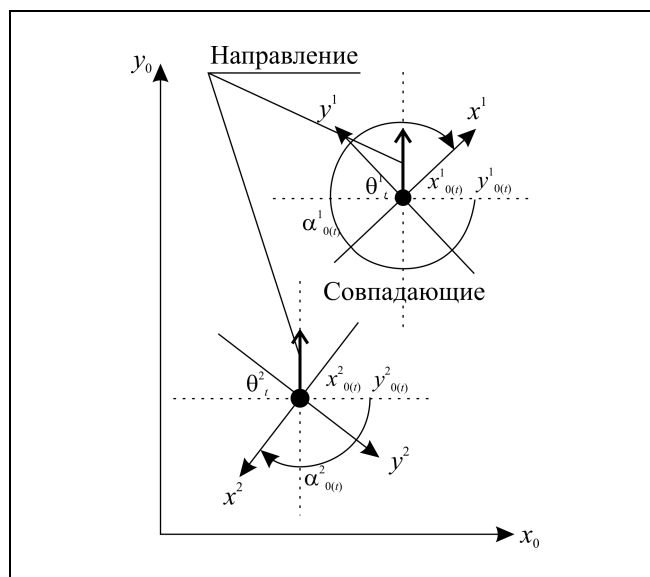


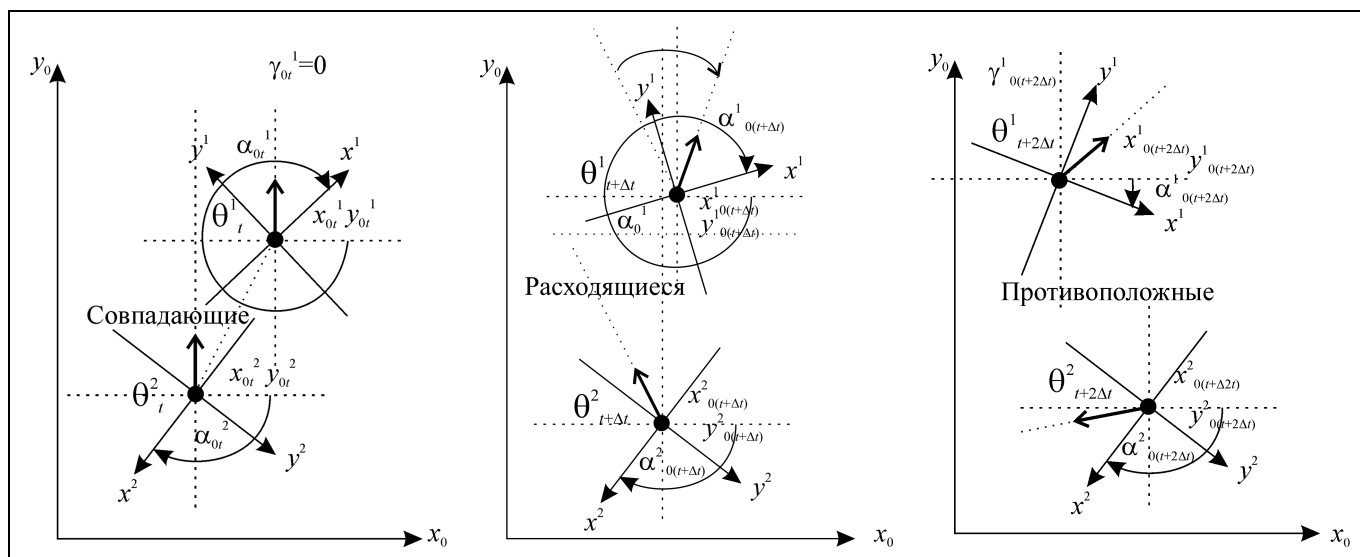
Рис. 2. Пример отношения на макросрезе: отношение направленный «совпадающие»

тров тяжести относительно оси x_0 . Объекты θ_t^1, θ_t^2 будем обозначать черными кружками, а их центры тяжести белыми кружками в центре объектов с координатами соответственно $x_{0(t)}^1, y_{0(t)}^1$; $x_{0(t)}^2, y_{0(t)}^2$ (рис. 2). Используя признаки нулевого уровня в качестве исходных, зададим отношения между объектами θ_t^1, θ_t^2 . Примерами таких отношений служат отношения расстояния, ориентации и направления движения.

Направления движения объектов изобразим вектором с одинарной стрелкой. Расстояние, как правило, является симметричным отношением и вычисляется как декартово расстояние между центрами тяжести. Для качественных рассуждений обычно вводится несколько качественных отношений расстояния, например, *маленькое*(θ_t^1, θ_t^2), *среднее*(θ_t^1, θ_t^2), *большое*(θ_t^1, θ_t^2).

Направление движения объектов друг относительно друга может быть задано как симметричным, так и несимметричным отношением. Аналогично отношению расстояния, введем несколько качественных отношений направления, например, симметричные отношения *совпадающие*(θ_t^1, θ_t^2), *расходящиеся*(θ_t^1, θ_t^2), *противоположные*(θ_t^1, θ_t^2). Пример отношения направления *совпадающие* приведен на рис. 2.

Отношение ориентации, как правило, не является симметричным, например, для качествен-


Рис. 3. Макротренд $Y_{напр}[\Theta_r, \Theta_t + 2\Delta t]$

ных рассуждений могут быть введены качественные отношения *справа*(θ_t^1, θ_t^2), *слева*(θ_t^1, θ_t^2), *сзади*(θ_t^1, θ_t^2), *спереди*(θ_t^1, θ_t^2), *справа*(θ_t^2, θ_t^1), *слева*(θ_t^2, θ_t^1), *сзади*(θ_t^2, θ_t^1), *спереди*(θ_t^2, θ_t^1).

Отношения на рис. 2 и 3 выделены жирным увеличенным шрифтом. Если эти отношения считать макроотсчетами, то их совокупность можно представить макросрезом $K(\Theta_r) = K(\theta_t^1, \theta_t^2) = \{\text{маленькое}(\theta_t^1, \theta_t^2), \text{совпадающие}(\theta_t^1, \theta_t^2), \text{сзади}(\theta_t^1, \theta_t^2), \text{спереди}(\theta_t^1, \theta_t^2), \text{справа}(\theta_t^1, \theta_t^2), \text{слева}(\theta_t^1, \theta_t^2), \text{сзади}(\theta_t^2, \theta_t^1), \text{спереди}(\theta_t^2, \theta_t^1), \text{справа}(\theta_t^2, \theta_t^1), \text{слева}(\theta_t^2, \theta_t^1)\}$. В общем случае отношения на макросрезе $K(\theta_t^1, \theta_t^2)$ между макроотсчетами могут быть различными и описываться на различных языках. Если все макроотсчеты данного макросреза рассматривать как предикаты, которые истинны или ложны, то отношение между ними может быть выражено формулами логики предикатов первого порядка. Например, если в момент времени t все предикаты *маленькое*(θ_t^1, θ_t^2), *совпадающие*(θ_t^1, θ_t^2), *сзади*(θ_t^1, θ_t^2), *спереди*(θ_t^1, θ_t^2) истинны, то это означает истинность формулы:

$$\begin{aligned} \Phi(\theta_t^1, \theta_t^2) = & \text{маленькое}(\theta_t^1, \theta_t^2) \wedge \\ & \wedge \text{совпадающие}(\theta_t^1, \theta_t^2) \wedge \text{сзади}(\theta_t^1, \theta_t^2) \wedge \\ & \wedge \text{спереди}(\theta_t^1, \theta_t^2). \end{aligned}$$

2.2. Отношения на макротрендах

На рис. 3 показаны отношения направления в моменты времени $t, t + \Delta t, t + 2\Delta t$ (на интервале $[t, t + 2\Delta t]$). Координаты центров тяжести, углы между осями на рисунках индексированы соответствующими моментами времени. Качественные признаки не индексированы. Используя введенные термины, можем говорить, что рис. 3 представляет макротренд:

$$Y_{напр}[\Theta_r, \Theta_t + 2\Delta t] = \langle \text{совпадающие}(\Theta_t), \text{расходящиеся}(\Theta_{t+\Delta t}), \text{противоположные}(\Theta_{t+2\Delta t}) \rangle,$$

где $\Theta_t = \{\theta_t^1, \theta_t^2\}, t \in [t, t + 2\Delta t]$.

Так же макротренды для отношения ориентации и расстояния могут быть записаны в виде:

$$\begin{aligned} Y_{ориент}[\Theta_r, \Theta_t + 2\Delta t] = & \langle \text{сзади}(\theta_t^1, \theta_t^2), \text{сзади}(\theta_t^2, \theta_t^1), \\ & \text{справа}(\theta_{t+\Delta t}^1, \theta_{t+\Delta t}^2), \text{справа}(\theta_{t+\Delta t}^2, \theta_{t+\Delta t}^1), \\ & \text{справа}(\theta_{t+2\Delta t}^1, \theta_{t+2\Delta t}^2), \text{справа}(\theta_{t+2\Delta t}^2, \theta_{t+2\Delta t}^1) \rangle, \\ Y_{расст}[\Theta_r, \Theta_t + 2\Delta t] = & \langle \text{маленькое}(\Theta_t), \text{среднее}(\Theta_{t+\Delta t}), \\ & \text{большое}(\Theta_{t+2\Delta t}) \rangle. \end{aligned}$$

Отношения на макротрендах временные. Если, как и для случая макросрезов, макроотсчеты данного макротренда рассматривать как предикаты, которые истинны или ложны, то отношение между ними также может быть выражено формулами логики предикатов первого порядка. Например, если полагать, что все предикаты *сзади*(θ_t^1, θ_t^2), *справа*(θ_t^1, θ_t^2), *спереди*(θ_t^1, θ_t^2), *слева*(θ_t^1, θ_t^2), *сзади*(θ_t^2, θ_t^1), *справа*(θ_t^2, θ_t^1), *спереди*(θ_t^2, θ_t^1), *слева*(θ_t^2, θ_t^1) истинны, то это означает истинность формулы:

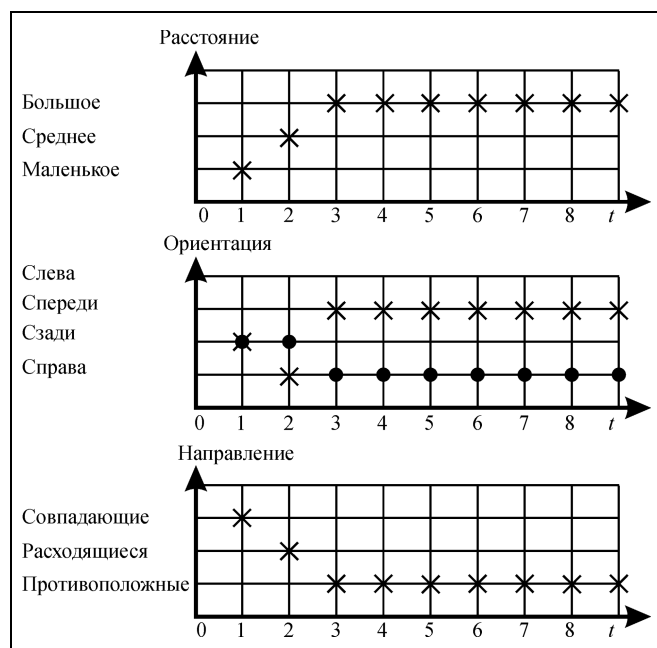


Рис. 4. Макротренды $Y_{\text{расст}}[\Theta_1, \Theta_9]$, $Y_{\text{ориент}}[\Theta_1, \Theta_9]$, $Y_{\text{напр}}[\Theta_1, \Theta_9]$

$ди(\theta_t^2, \theta_t^1)$, $справа(\theta_{t+\Delta t}^1, \theta_{t+\Delta t}^2)$, $сзади(\theta_{t+\Delta t}^2, \theta_{t+\Delta t}^1)$, $справа(\theta_{t+2\Delta t}^1, \theta_{t+2\Delta t}^2)$, $справа(\theta_{t+2\Delta t}^2, \theta_{t+2\Delta t}^1)$, соответствующие макротренду $Y_{\text{ориент}}[\Theta_p, \Theta_{t+2\Delta t}]$, истинны в соответствующие моменты времени, то это означает истинность, например, следующей формулы:

$$\begin{aligned} \Phi_{\text{ориент}}[\Theta_p, \Theta_{t+2\Delta t}] = & [((сзади(\theta_t^1, \theta_t^2) \wedge \\ & \wedge сзади(\theta_t^2, \theta_t^1)) \supset ((справа(\theta_{t+\Delta t}^1, \theta_{t+\Delta t}^2) \wedge \\ & \wedge сзади(\theta_{t+\Delta t}^2, \theta_{t+\Delta t}^1))) \supset (справа(\theta_{t+2\Delta t}^1, \theta_{t+2\Delta t}^2) \wedge \\ & \wedge справа(\theta_{t+2\Delta t}^2, \theta_{t+2\Delta t}^1)). \end{aligned}$$

Существует немало подходов к формулировке и вычислению причинно-следственных отношений во времени. Один из них основан на отношениях временных интервалов. В соответствии с работой [6] существует 13 бинарных временных отношений между двумя временными интервалами, заданными на дискретной шкале времени.

Каждый тренд определен на временном интервале, и поэтому указанные отношения могут использоваться и для трендов. Тренды содержат значения признаков, соответствующих определенным моментам времени. Например, рассмотренные нами тренды можно представить в виде, показанном

на рис. 4 (крестикам и точкам в узлах сетки на оси ординат соответствует отсчет, а на оси абсцисс — момент времени, соответствующий этому отсчету).

2.3. Отношения на макросрезах и макротрендах

Разобьем интервал $[\Theta_1, \Theta_9]$ (см. рис. 4) на подинтервалы с одинаковыми наборами значений макросрезов на всех трендах. В результате получим рис. 5, где на каждом интервале задан поток одинаковых макросрезов, изображенных прямоугольниками.

В результате выделения таких потоков макросрезов совокупность макротрендов $Y_{\text{расст}}[\Theta_1, \Theta_9]$, $Y_{\text{ориент}}[\Theta_1, \Theta_9]$, $Y_{\text{напр}}[\Theta_1, \Theta_9]$ можно представить как объединение потоков макросрезов, т. е.

$$\begin{aligned} Y_{\text{расст, ориент, напр}}[\Theta_1, \Theta_9] = & Y_{\text{расст, ориент, напр}}[\Theta_1, \Theta_1] \cup \\ \cup & Y_{\text{расст, ориент, напр}}[\Theta_2, \Theta_2] \cup Y_{\text{расст, ориент, напр}}[\Theta_3, \Theta_9]. \end{aligned}$$

Разбиение макротрендов на интервалы с одинаковыми макросрезами позволяет представить отношение $Y_{\text{расст, ориент, напр}}[\Theta_1, \Theta_9]$ для макротренда с помощью интервальных отношений *перед* и *предшествует*. Отношения на макросрезах и макротрендах могут быть описаны на различных языках. В случае достаточно полных знаний об этих отношениях во времени и пространстве удобно воспользоваться конечно-автоматными моделями.

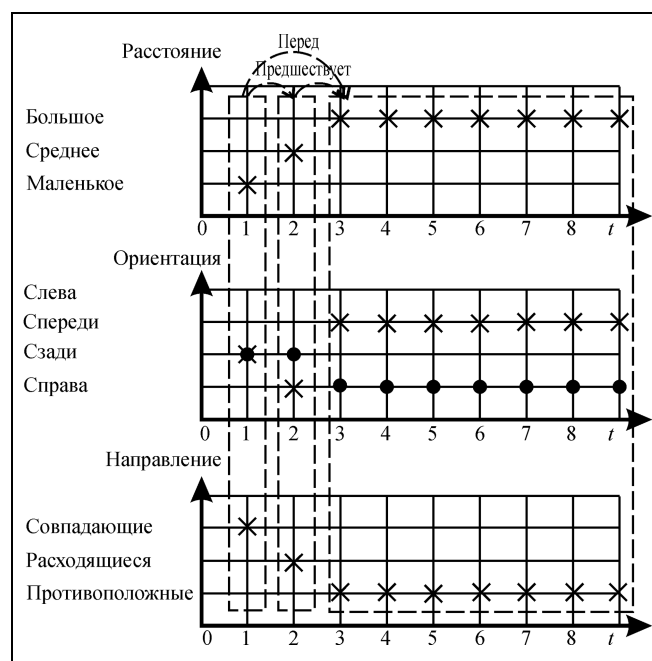


Рис. 5. Одинаковые макросрезы на макротрендах $Y_{\text{расст, ориент, напр}}[\Theta_1, \Theta_9]$

3. ЗАДАНИЕ ОТНОШЕНИЙ КОНЕЧНО-АВТОМАТНЫМИ МОДЕЛЯМИ

3.1. Иерархическая модель четких конечных автоматов

Представим макротренд $Y_i[\Theta_{t_s}, \Theta_{t_e}] = \langle y_i(\Theta_{t_s}), \dots, y_i(\Theta_{t_e}) \rangle$ конечным автоматом $M_i[\Theta_{t_s}, \Theta_{t_e}]$ следующим образом. Начальным внутренним состоянием автомата будем считать состояние $b_i(\Theta_{t_s})$. Функциями переходов f и выходов ϕ конечного автомата $M_i[\Theta_{t_s}, \Theta_{t_e}]$ будем считать следующие функции:

$$f[t, b_i(\Theta_t)] = b_i(\Theta_{t+\Delta}), t \in [t_s, t_e - \Delta],$$

$$\phi[b_i(\Theta_t)] = y_i(\Theta_t).$$

Таким образом, вся информация о трендах и макротрендах, срезах и макросрезах может быть представлена иерархией конечных автоматов, начиная от автоматов, представляющих тренды признаков нулевого уровня и заканчивая автоматами, представляющими макротренды признаков самого верхнего уровня. Автоматы, представляющие признаки уровня u , будем называть автоматами уровня u . В случае необходимости указания уровня автомата $M_i[\Theta_{t_s}, \Theta_{t_e}]$ будем автомат записывать

как $M_i^u[\Theta_{t_s}, \Theta_{t_e}]$. Если представить автоматы графами переходов, то для нашего примера будем иметь два уровня автоматов, задающих качественные отношения: $M_{\text{расст}}^1[\Theta_{t_s}, \Theta_{t_e}]$, $M_{\text{ориент}}^1[\Theta_{t_s}, \Theta_{t_e}]$,

$M_{\text{напр}}^1[\Theta_{t_s}, \Theta_{t_e}]$ и $M_{\text{расст, ориент, напр}}^2[\Theta_{t_s}, \Theta_{t_e}]$. Построение автоматов начинается с нулевого уровня по трендам отсчетов нулевого уровня. Каждый автомат некоторого уровня $u > 0$ представляет динамику изменения отношений между отношениями, представляемыми автоматами предыдущих уровней. Будем полагать, что все автоматы $M_{k_u}^u$ одного и того же уровня проиндексированы от 1 до l_u , т. е. каждый автомат $M_{k_u}^u$ имеет уникальный номер $k_u \in \{1, \dots, l_u\}$.

Источником трендов служит человек, участвующий в мультимодальном интерфейсе и либо реальная физическая среда (мир), либо ментальный (воображаемый) мир, либо их комбинация. Поведение человека и миров может описываться и непосредственно четкими конечными автоматами,

минуя этап выделения отдельных трендов. Эти вопросы выходят за рамки настоящей статьи, тем более что им посвящено огромное число работ по конечным автоматам.

Сущность метода распознавания отношений, использующего четкие автоматы, состоит в следующем. Построим заранее все необходимые четкие автоматы $M_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$, $k_u = 1, \dots, l_u$ по трендам $Y_{i_u}[\Theta_{t_s}, \Theta_{t_e}] = \langle y_{i_u}(\Theta_{t_s}), \dots, y_{i_u}(\Theta_{t_e}) \rangle$, $i_u \in \{1, \dots, m_u\}$, задающим подлежащие распознаванию пространственно-временные отношения в среде между совокупностями объектов Θ , называемых *эталонными*. Эталонные объекты являются образами объектов, отношения между которыми будут впоследствии распознаваться. Эти объекты могут формироваться экспертно (ментально) путем моделирования поведения пользователей и среды или путем априорных наблюдений за ними. Для распознавания

отношений установим все автоматы $M_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$, $k_u = 1, \dots, l_u$, в начальное состояние, соответствующее моменту времени t_s , и начнем подавать на них синхронно временные метки (моменты времени) интервала $[t_s, t_e]$, вычисляя с помощью функций переходов и выходов $f[t, b_i(\Theta_t)] = b_i(\Theta_{t+\Delta})$, $t \in [t_s, t_e - \Delta]$, $\phi[b_i(\Theta_t)] = y_i(\Theta_t)$ отношения $y_i(\Theta_t)$ и сравнивая их с наблюдаемыми отношениями $y_i(\hat{\Theta}_t)$, где $\hat{\Theta}_t$ — наблюдаемые объекты, каждому из которых взаимно однозначно соответствует эталонный объект множества Θ_t . Если окажется, что

для всех $t \in [t_s, t_e]$ имеет место $y_i(\Theta_t) = y_i(\hat{\Theta}_t)$, то считается, что пространственно-временное отношение между объектами кортежа $\langle \hat{\Theta}_{t_s}, \dots, \hat{\Theta}_{t_e} \rangle$, задаваемое автоматами $M_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$, $k_u = 1, \dots, l_u$, распознано.

Если ввести характеристическую функцию

$$\chi(y_i(\Theta_t), y_i(\hat{\Theta}_t)) = \begin{cases} 0, & \text{если } y_i(\Theta_t) \neq y_i(\hat{\Theta}_t), \\ 1, & \text{если } y_i(\Theta_t) = y_i(\hat{\Theta}_t), \end{cases}$$

то пространственно-временное отношение между объектами кортежа $\langle \hat{\Theta}_{t_s}, \dots, \hat{\Theta}_{t_e} \rangle$, задаваемое автоматом $M_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$, $k_u = 1, \dots, l_u$, если распознано,

$\chi(y_i(\Theta_t), y_i(\hat{\Theta}_t)) = 1$ для всех $t \in [t_s, t_e]$. В противном случае значение характеристической функции будет равно нулю хотя бы для одного $t \in [t_s, t_e]$.



3.2. Иерархическая модель нечетких конечных автоматов

Фаззифицируем автомат $M_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$ следующим образом. Каждому отношению $y_i(\Theta_t)$ сопоставим упорядоченное множество $\tilde{Y}_i(\Theta_t)$ отношений $\tilde{y}_i(\Theta_t)$, характеризующих совокупность объектов Θ в момент времени t . Зададим функцию принадлежности $\mu_{\tilde{Y}_i} = \mu_{\tilde{Y}_i}[\tilde{y}_i(\Theta_t)]$ на множестве $\tilde{Y}_i(\Theta_t)$. В целях снижения вычислительной сложности для интеллектуального мультимодального интерфейса наиболее подходят треугольные функции принадлежности. Пару $(\mu_{\tilde{Y}_i}[\tilde{y}_i(\Theta_t)], \tilde{Y}_i(\Theta_t))$ будем называть *нечетким отношением*. Функция принадлежности задается экспертно или получается в результате обработки экспериментальных данных. Последовательность $\tilde{Y}_i[\Theta_{t_s}, \Theta_{t_e}] = \langle \tilde{Y}_i(\Theta_{t_s}), \dots, \tilde{Y}_i(\Theta_{t_e}) \rangle$ будем называть *нечетким макротрендом*. Функциями переходов $\tilde{f}$ и выходов $\tilde{\phi}$ нечеткого конечного автомата $\tilde{M}_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$ будем считать следующие функции:

$$\tilde{f}[t, b_i(\Theta_t)] = b_i(\Theta_{t+\Delta}), \quad t \in [t_s + \Delta, t_e - \Delta],$$

$$\tilde{f}[t_s] = b_i(\Theta_{t_s}),$$

$$\tilde{\phi}[b_i(\Theta_t)] = [\mu_{\tilde{Y}_i}(\tilde{y}_i(\Theta_t)), \tilde{Y}_i(\Theta_t)].$$

Четкий автомат является частным случаем нечеткого и отличается от нечеткого только функцией выходов. В случае четкого автомата функция выходов возвращает значение макроотсчета, а в случае нечеткого — нечеткое отношение (множество макроотсчетов и функцию принадлежности, заданную на нем). Сущность метода распознавания отношений, использующего нечеткие конечные автоматы, состоит в следующем. Построим заранее все необходимые нечеткие автоматы $\tilde{M}_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$ по эталонным трендам $\tilde{Y}_i[\Theta_{t_s}, \Theta_{t_e}] = \langle \tilde{Y}_i(\Theta_{t_0}), \dots, \tilde{Y}_i(\Theta_{t_e}) \rangle$, задающим подлежащие распознаванию пространственно-временные нечеткие эталонные отношения между совокупностями объектов Θ . Установим все автоматы в начальное состояние, соответствующее моменту времени t_s , и начнем подавать на все автоматы $\tilde{M}_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$ синхронно временные метки (моменты времени) интервала $[t_s, t_e]$, вычисляя с по-

мощью функций переходов и выходов нечеткие отношения $(\mu_{\tilde{Y}_i}[\tilde{y}_i(\Theta_t)], \tilde{Y}_i(\Theta_t))$. Если окажется, что для всех $t \in [t_s, t_e]$ отсчет $\tilde{y}_i(\hat{\Theta}_t) \in \tilde{Y}_i(\hat{\Theta}_t)$, где $\hat{\Theta}_t$ — наблюдаемые отношения такие, что $\hat{\Theta}_t = \Theta_t$, то считается, что пространственно-временное отношение между отношениями кортежа $\langle \hat{\Theta}_{t_s}, \dots, \hat{\Theta}_{t_e} \rangle$, задаваемое автоматом $\tilde{M}_{k_u}^u[\Theta_{t_s}, \Theta_{t_e}]$, распознано. Значение функции принадлежности $\mu_{\tilde{Y}_i}[\tilde{y}_i(\hat{\Theta}_t)]$ характеризует степень принадлежности отчета $\tilde{y}_i(\hat{\Theta}_t)$ множеству отсчетов $\tilde{Y}_i(\Theta_t)$.

4. ПРИНЦИП АГРЕГИРОВАНИЯ МУЛЬТИМОДАЛЬНОГО ИНТЕРФЕЙСА И РАСПОЗНАВАНИЯ ДИНАМИЧЕСКИХ СИТУАЦИЙ

Любой автомат $\tilde{M}_{k_u}^u$ уровня $u \geq 1$ задает пространственно-временные отношения между отношениями уровня $u - 1$. Таким образом, уровни автоматов используются для представления иерархии отношений. Автомат $\tilde{M}_{k_u}^u$ любого уровня, отношения которого не используются для задания отношений более высокого уровня автоматами более высокого уровня, назовем *финальным*. Кликкой $\tilde{K}_{k_u}^u$ финального автомата $\tilde{M}_{k_u}^u$ назовем множество подмножеств нечетких автоматов $K_{k_u}^u, \dots, K_{k_0}^0, k_u \in \{1, \dots, l_u\}$ уровней $u \geq 0$, получаемое, начиная с финального автомата $\tilde{M}_{k_u}^u$, в соответствии со следующим алгоритмом.

Шаг 1. Принимаем $K_{k_u}^u = \{\tilde{M}_{k_u}^u\}$. Если $u = 0$, то принимаем $\tilde{K}_{k_u}^u = K_{k_u}^u$ и переходим к шагу 3. В противном случае принимаем $v = u$. Принимаем $v = v - 1$ и переходим к шагу 2.

Шаг 2. Для каждого автомата $\tilde{M}_{k_v}^v \in K_v^v$ получаем множество $K_{k_{v-1}}^{v-1}$ автоматов, используемых для представления отношений, участвующих в построении этого автомата $\tilde{M}_{k_v}^v \in K_v^v$, т. е. $K_{k_{v-1}}^{v-1} = \{\tilde{M}_{k_{v-1}}^{v-1} | \text{отношения, задаваемые автоматом } \tilde{M}_{k_{v-1}}^{v-1} \text{ участвуют в построении } \tilde{M}_{k_v}^v\}$. Если $v - 1 = 0$, то

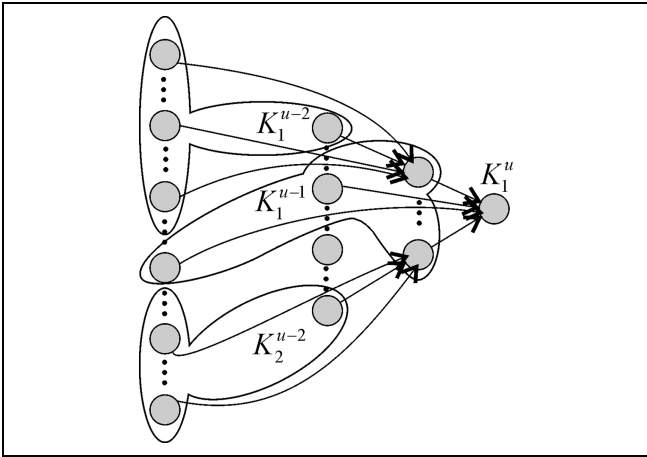


Рис. 6. Пример клики $\tilde{K}^u$ финального автомата $\tilde{M}^u$

$\tilde{K}_{k_u}^u = K_{k_{u-1}}^{u-1} \cup K_{k_{u-2}}^{u-2} \cup \dots \cup K_{k_0}^0$ и переходим к шагу 3. В противном случае принимаем $v = v - 1$ и переходим к шагу 2.

Шаг 3. Конец построения клики $\tilde{K}_{k_u}^u$. ♦

Пример клики $\tilde{K}^u$ финального автомата $\tilde{M}^u$ показан на рис. 6, где кружки изображают автоматы. Стрелки направлены от автоматов более низкого уровня к автоматам более высокого уровня, отношения которых строятся на основе отношений, представляемых автоматами, соединенными с ними входящими стрелками.

Каждая клика $\tilde{K}_{k_u}^u$, $k_u \in \{1, \dots, l_u\}$ задает совокупные динамические отношения объектов пользователя и объектов среды в процессе взаимодействия через интеллектуальный интерфейс, направленного на распознавание той или иной динамической ситуации. Эта клика создается заранее путем наблюдения за пользователем и средой или экспертно. Например, пользователь может в процессе взаимодействия пользоваться тактильными модальностями, т. е. прикасаться к определенным местам экрана навигатора автомобиля, являющимся по отношению к пользователю внешней средой, в целях получения интересующей его информации. Тогда та или иная клика должна содержать совокупные динамические отношения между водителем и навигатором, позволяющие распознать, например, автора прикосновения, порядок и координаты прикосновения, координаты автомобиля, обстановку вокруг него и т. п.

Задача интеллектуального мультимодального интерфейса в процессе пространственно-временного распознавания динамических ситуаций может рассматриваться как задача вычисления отношений на кликах, т. е. в общем случае надо вы-

числить, какая или какие клики соответствуют пространственно-временным отношениям (ситуациям) между пользовательскими объектами и объектами среды. Пользователь является источником макротрендов, задающих его поведение в процессе интеллектуального интерфейса.

Сущность метода распознавания динамических ситуаций на кликах состоит в следующем. Имеем множество отношений $\{\Theta_t^{k_u} | t \in [t_s, t_e], k_u \in \{1, \dots, l_u\}, u \in \{1, \dots, R\}\}$ и множество клик $\tilde{K}_{k_u}^u$ (здесь R — максимальное число уровней), построенное по множеству потоков макротрендов $\mathcal{U}_u = \{\tilde{Y}_{k_u}^{k_u}[\Theta_{t_s}^{k_u}, \Theta_{t_e}^{k_u}] | k_u \in \{1, \dots, l_u\}, u \in \{1, \dots, R\}\}$, где $\tilde{Y}_{k_u}^{k_u}[\Theta_{t_s}^{k_u}, \Theta_{t_e}^{k_u}] = \langle \tilde{y}_{k_u}(\Theta_{t_s}^{k_u}), \dots, \tilde{y}_{k_u}(\Theta_{t_e}^{k_u}) \rangle$.

Шаг 1. Установить $t = t_s$.

Шаг 2. Наблюдая за поведением пользователя и среды вычислить для всех наблюдаемых отношений $\hat{\Theta}_t^{k_u}$, $k_u \in \{1, \dots, l_u\}$, $u \in \{0, \dots, R\}$ макросрезы $\{\tilde{y}(\hat{\Theta}_t^1), \dots, \tilde{y}(\hat{\Theta}_t^{l_u})\}$. Подать на каждый автомат $\tilde{M}_{k_u}^u$, $k_u = 1, \dots, l_u$, $u = 0, \dots, R$ входное воздействие t ; вычислить, используя функции переходов, состояния $b(\Theta_t^{k_u})$ и значения функций выходов $\tilde{\phi}[b(\Theta_t^{k_u})] = [\mu_{\tilde{Y}}(\tilde{y}(\Theta_t^{k_u})), \tilde{Y}(\Theta_t^{k_u})]$.

Шаг 3. Если $t = t_e$, то перейти к шагу 4. В противном случае установить $t = t + \Delta t$ и перейти к шагу 2.

Шаг 4. Если окажется, что существуют финальные автоматы $\tilde{M}_{k_u}^u$ такие, что для всех $t \in [t_s, t_e]$ и всех автоматов клики $\tilde{K}_{k_u}^u$ $\tilde{y}_i(\hat{\Theta}_t) \in \tilde{Y}_i(\Theta_t)$, где $\hat{\Theta}_t$ — наблюдаемые отношения такие, что $\hat{\Theta}_t = \Theta_t$, то считается, что пространственно-временное отношение между отношениями кортежей $\langle \hat{\Theta}_{t_s}, \dots, \hat{\Theta}_{t_e} \rangle$, задаваемое автоматами клики $\tilde{K}_{k_u}^u$ на интервале $[t_s, t_e]$, распознано.

Шаг 5. Вычислить степень уверенности [15] результатов распознавания ситуации с помощью клики $\tilde{K}_{k_u}^u$, используя один из операторов агрегирования (например, оператор Шоке или Сугено [16]), значений функций принадлежности $\mu_{\tilde{Y}_i}[\tilde{Y}_i(\hat{\Theta}_t)]$ для всех $t \in [t_s, t_e]$ и всех автоматов клики $\tilde{K}_{k_u}^u$. ♦



ЗАКЛЮЧЕНИЕ

Предложен метод организации интеллектуального мультимодального интерфейса пользователя со средой в процессе распознавания возникающих в среде динамических ситуаций на основе совокупности взаимодействующих нечетких конечных автоматов. Он может применяться в различных областях, требует сравнительно мало ресурсов по времени и памяти, позволяет достаточно естественно, а главное, без длительного периода обучения, формировать иерархию нечетких пространственно-временных отношений.

Предлагаемый подход реализован в аудио-визуальном интерфейсе «Human — TV set» (разработка в МГТУ им. Н.Э. Баумана для фирмы NXP Semiconductors founded by Philips), который продемонстрировал надежную работу. В аппаратной части интерфейса используются две VGA камеры, один ненаправленный РС микروفон, мультимедийная вычислительная платформа STB810 с портированными алгоритмами захвата и отслеживания объекта. Распознавание пользователя, жестовых и голосовых команд происходит в реальном времени (менее 1 с) в помещении при обычном электрическом или дневном освещении, без специальных маркеров на теле пользователя.

ЛИТЕРАТУРА

1. Раскин Д. Интерфейс: новые направления в проектировании компьютерных систем. — М: Символ-Плюс, 2005. — С. 120—129.
2. Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction / Z. Duric, et al. // IEEE(90). — 2002. — N 7. — P. 1272—1289.
3. Intelligent user interface based on multimodal dialog control for audio-visual systems / H. Shinjo, et al. // Hitachi Hyoron. — 2005. — Vol. 87, N 10. — P. 801—804.
4. Brooks R. A. Cambrian Intelligence. — Cambridge, MA: MIT Press, 1999. — 225 p.

5. Clementini E., Felice P.D., Hernandez D. Qualitative representation of positional information // Artificial Intelligence. — 1997. — 95(2). — P. 317—356.
6. Allen F. Maintaining knowledge about temporal intervals // Communications of the ACM. — 1983. — 26(11). — P. 832—843.
7. Cohn A.G., Hazarika S. M. Qualitative spatial representation and reasoning: An overview // Fundamenta Informaticae. — 2001. — 46(1). — P. 1—29.
8. Representing and reasoning with qualitative spatial relations about regions / A.G. Cohn, et al. // In Spatial and Temporal Reasoning. — 1997. — P. 97—134.
9. Freksa C. Using orientation information for qualitative spatial reasoning // Proc. of Intern. Conf. GIS: Theories and Methods of Spatio-Temporal Reasoning. — 1992. — Vol. 639. — P. 162—178.
10. Spar Q: A toolbox for qualitative spatial representation and reasoning / F. Dylla, et al. // In Qualitative Constraint Calculi: Application and Integration. — 2006. — P. 79—90.
11. Musto A. On spatial reference frames in qualitative motion representation // Technical Report FKI-230-99, Institut fur Informatik, Technische Universitat Munchen. — 1999. — 18 p.
12. A qualitative trajectory calculus as a basis for representing moving objects in geographical information systems / N. Van de Weghe, et al. // Control and Cybernetics. — 2006. — Vol. 35. — P. 97—119.
13. Holzmann C. Rule-Based Reasoning about Qualitative Spatio-temporal Relations // MPAC2007. — 2007. — P. 49—54.
14. Поспелов Д.А. Интеллектуальные интерфейсы для ЭВМ новых поколений // Электронная вычислительная техника. — 1989. — № 3. — С. 4—20.
15. Девятков В.В., Алфимцев А.Н. Распознавание манипулятивных жестов // Вестник МГТУ им. Н.Э. Баумана / Сер. Приборостроение. — 2007. — № 3. — С. 56—75.
16. Devyatkov V., Alfimtsev A. Optimal Fuzzy Aggregation of Secondary Attributes in Recognition Problems // Proc. of 16th Intern. Conf. in Central Europe on Computer Graphics, Visualization and Computer Vision. — Plzen, 2008. — P. 78—85.

Статья представлена к публикации членом редколлегии А.А. Амбарцумяном.

Девятков Владимир Валентинович — д-р техн. наук, зав. кафедрой, ☎ (499) 267-65-37, ✉ deviatkov@iu3.bmstu.ru,

Алфимцев Александр Николаевич — канд. техн. наук, доцент, ☎ (499) 263-66-25, ✉ alfim@iu3.bmstu.ru,

Московский государственный технический университет им. Н.Э. Баумана.

Новая книга

Сетевые модели в управлении / Сб. статей под ред. Д.А. Новикова, О.П. Кузнецова, М.В. Губко. — М.: Эгвес, 2011. — 443 с.

Настоящее издание сформировано на основе материалов специального выпуска электронного Сборника трудов «Управление большими системами» (ubs.mtas.ru) и посвящено задачам управления, в которых объект управления (и (или) система управления) имеет сетевую структуру.

Статьи разбиты по рубрикам, отражающим скорее неформальную группировку по актуальным научным направлениям, чем строгую, претендующую на полноту, классификацию:

- сетцентрическое управление и многоагентные системы;
- управление технологическими сетями;
- сетевые модели в принятии решений;
- когнитивные карты в управлении;
- сетевые организации и социальные сети.

Можно надеяться, что настоящий сборник, демонстрируя единство возможных подходов к решению задач сетевого управления объектами самой разной природы, не только будет интересен для ученых и практиков, но и сможет дать почву для интеграции усилий специалистов в разных разделах теории управления.