

МОДЕЛИРОВАНИЕ ПРОЦЕССА ДИАЛОГОВОГО ОБУЧЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ САМООРГАНИЗУЮЩИХСЯ СИСТЕМ

Авторами разработаны методологические принципы с доведением до конкретных структурных модельных схем: создание модели, средствами которой можно воспроизвести динамику процесса обучения ИСС, исходя из его структуры, безотносительно к конкретному содержанию.

Ключевые слова: модель, интеллектуальная самоорганизующаяся система.

Введение

В современном технологическом мире *обучение* есть ключевой элемент формирования специалиста. Общеизвестен (отмечался классиками ещё в 16-м – 17-м веках) факт лавинообразного нарастания массива человеческих знаний. Приблизённо, к концу 19-го века технические знания ещё «оставались актуальными» на протяжении 1–2 человеческих жизней. К середине 20-го века темпы обновления стали соизмеримы с продолжительностью индивидуальной человеческой жизни, а к концу века – превысили её в 2–3 раза. Эта тенденция достаточно наглядна даже в бытовой сфере, на примере развития компьютерной техники: 10–15 лет – время фатального устаревания не только аппаратного и программного обеспечения, но и ведущих используемых физических (физико-технологических) принципов, заложенных в их основу. В связи с указанным, адекватно должны возрастать темпы и качество подготовки и переподготовки специалистов по соответствующим научно-техническим направлениям. В этом аспекте актуальны исследования по *диалоговым системам обучения* (ДСО). Представляют интерес, в частности:

- разработка методов синтеза, наполнения (адаптации), сопровождения, аттестации и проч. ДСО;
- разработка эффективных методов подачи и контроля предметного (прикладного, содержательного) материала в ДСО;
- сравнение и оценка эффективности конкретных имеющихся ДСО;
- разработка методов модифицирования ДСО в процессе их сопровождения, в направлении повышения их эффективности.

В связи с перечисленными и рядом других подобных направлений, *актуально* применение методов моделирования. Обучение – как процесс сооб-

щения знаний индивидуальным носителям *человеческого интеллекта* (ЧИ) – происходит во времени. Учащиеся (субъекты процесса обучения) присутствуют в этом процессе «во множественном числе». Легко видеть, что представленная ситуация в целом соответствует парадигме параллельного распределённого моделирования.

Другая область, в которой применимы сходные подходы, – обучение интеллектуальных *компьютеризированных систем* (КС). Назначение КС – быть *усилителями человеческого интеллекта* (УЧИ) [1]. Этот же ЧИ неизменно (в той или иной степени) является прототипом КС, чем обусловлено единство путей развития ЧИ и интеллекта КС [2]. Другие, *принципиально альтернативные* прототипы – попросту отсутствуют, либо не наблюдаются (не воспринимаются) человеком, как носители (обладатели) другого особого *неантропоморфного* интеллекта. Так, в частности, интеллект животных, как объект исследования, есть по существу проекция человеческих представлений об интеллекте, т.е. проекция ЧИ.

УЧИ способны быть таковыми только если они «максимально подобны» ЧИ. В противном случае их *неантропоморфный* интеллект будет отторгнут (не понят, не воспринят) ЧИ. На современном этапе этот «принцип максимального подобия» реализуется, в частности, применением дружественных и интеллектуальных интерфейсов в КС. В дальнейшем, по видимому, человечество ожидает глобальная роботизация и глобальная же интеллектуализация всего человеческого окружения. Таким образом, интеллектуальные КС перспективны именно в смысле реализации в них *интеллектуальных самообучающихся систем* (ИСС) – аналога ЧИ, а, следовательно, с использованием в той или иной форме принципов ДСО. На текущий момент данный процесс развёртывается, в частности с использованием нейронных сетей, в частности на *самоорганизующихся картах Кохонена* (СКК) [3].

Цель настоящей работы – разработка методологических принципов с доведением до конкретных структурных модельных схем: создание модели, средствами которой можно воспроизвести динамику процесса обучения ИСС, исходя из его структуры, безотносительно к конкретному содержанию.

1. Понятия и определения. Концепция модели

Система *человеческих научных знаний* (ЧНЗ) – факты, обобщения, гипотезы и др. – весьма разнообразна: охватывает различные области взаимодействия человека с окружающим миром. Отдельные области ЧНЗ зачастую не имеют чётко очерченных границ, бывают разработаны (детализированы) в разной степени, нередко содержат пробелы и противоречия. Однако целом – система ЧНЗ в достаточной степени взаимосогласованная и структурирована, чтобы быть основой для развития технологий и др. Суммарный информационный объём ЧНЗ значительно превышает возможности хранения информации в одном экземпляре человеческого мозга. Во всём человечестве, как биологическом виде, не существует ни одной особи, которая «знала бы всё». Вся совокупность ЧНЗ хранится *вне* конкретных отдельных человеческих мозгов и представляет собой распределённую информационную систему.

В процессе развития человеческой цивилизации, отдельные элементы ЧНЗ передаются от одних человеческих особей другим посредством системы обучения. Ключевым понятием процесса обучения является *учебная дисциплина* (УД) – изучение конкретного раздела научного знания. В системном плане, УД характеризуется структурой и содержанием (рис. 1).

Содержательный план УД может быть достаточно разнообразным. Конкретный набор объектов содержательного плана определяется структурой представления научного знания, традиционно принятой в соответствующей ветви науки. УД может включать в себя: определения (формулировки); правила; алгоритмы (последовательности действий); навыки применения (методологические наработки, касающиеся формирования навыков) и др.

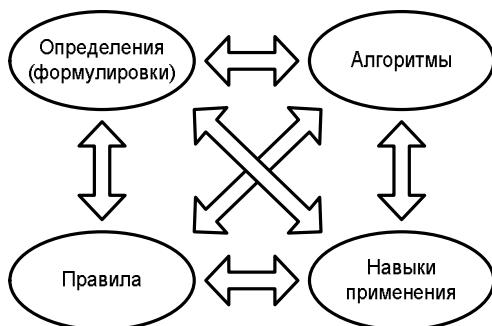


Рис. 1. Структура и содержание учебной дисциплины

Структурный план УД предполагает совокупность связей между указанными элементами содержательного плана. Характер связей и их число могут быть различными в зависимости от конкретного заполнения элементов содержательного плана. Степень взаимосвязанности элементов содержательного плана косвенно указывает на уровень проработанности научного знания, положенного в основу соответствующей УД.

На рис. 1 *содержательный план* показан в виде отдельных обобщённых областей, а *структурный план* – в виде набора двунаправленных прозрачных стрелок. Связи даны по принципу «каждый с каждым», поскольку представленный пример подсистемы содержит достаточно общие понятия.

Структурированностью УД, как конкретной ветви ЧНЗ, определяется потенциальная возможность построения модели процесса обучения. Простейший вариант – модель *диалоговой системы обучения* (ДСО) рис. 2. – воспроизводит диалог между «преподавателем» и «студентом», в ходе которого осуществляется передача знаний.

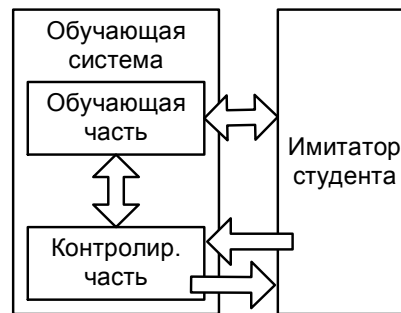


Рис. 2. Модель диалоговой системы обучения

Важная особенность модели – воспроизведение динамики запоминания результатов обучения. Обучающий диалог происходит не непрерывно, а в виде последовательности сеансов (уроков). Между сеансами имеются перерывы. За время перерыва происходит *забывание* – утрата части выученного знания. Включение перерывов в процесс обучения – не прихоть модели, а одна из важных особенностей реального процесса обучения. *Забывание* имеет место как в «живых системах» (как «человеческий фактор»), так и в случаях когда объектом обучения является КС («растренировка» при прогрессирующем накоплении знаний). Изучение (воспроизведение) динамики забывания в процессе обучения – есть одно из назначений модели.

Представляют интерес, в частности, различные механизмы забывания, связанные с различными видами функции забывания от времени. Интересны аддитивный, мультипликативный и степенной законы забывания:

- аддитивное забывание - $H(t_1) = H(t_0)(1 - (t_1 - t_0)k_a)$; $k_a < 1$ – аддитивный коэффициент забывания;

- мультипликативное забывание - $H(t_1) = H(t_0) \cdot k_m$; $k_m < 1$ – мультипликативный коэффициент забывания;

- степенное забывание - $H(t_1) = H(t_0)^{(1-k_p)k_p}$; $k_p < 1$ – степенной коэффициент забывания.

Предположительно, они могут быть сопоставлены с разными психологическими состояниями в процессе обучения. Варьирование психологическими состояниями (в модельной интерпретации) подбором надлежащих объёмов и темпа подачи обучающего материала – чрезвычайно интересный объект модельного исследования, допускающий непосредственное практическое использование.

На рис. 2 представлена схема модели ДСО, включающая «имитатор преподавателя» (*обучающая система*) и «имитатор студента». Обучающая система содержит обучающую и контролируемую часть. Обучающая часть проводит «аналог лекционных занятий», контролирующая часть – «аналог зачёта». Т.о., имеются две взаимосвязанные относительно автономные системы и, соответственно, воспроизводятся две фазы взаимодействия *обучающей системы* со «студентом»:

- диалог получения обучающей информации. Включая, возможно, интерфейс управления *обучающей частью обучающей системы*;

- диалог контроля знаний – *контролирующая часть обучающей системы*.

Типовая последовательность обучения, согласно модели, может быть следующей:

- задаётся (изначально устанавливается) выборка вопросов;

- вопросы извлекаются из выборки «с возвращением» и многократно предъявляются «имитатору студента»;

- каждый вопрос имеет сложность и «уровень изученности». Чем выше сложность, тем ниже вероятность правильного ответа;

- при нахождении правильного ответа – сложность вопроса снижается, а «уровень изученности» растёт;

- при снижении сложности вопроса до 0, или достижении «порогового уровня изученности» – вопрос считается выученным и ставится «не контроль»;

Модель может поддерживать многочисленные влияющие факторы: сложность вопроса; уровень изученности вопроса; коэффициент забывания; коэффициент усталости; регулярность занятий; пороговое число правильных ответов; пороговый показатель успешности студента – число освоенных элементов информации; пороговый показатель успешности курса – число успешно освоивших его студентов; пороговый индивидуальный и «средне-групповая» степень усвоения курса и др.

Представленной структурой поддерживается ситуация, в целом соответствует парадигме парал-

лельного распределённого моделирования. Обучение – есть процесс, происходящий во времени, а учащиеся (субъекты процесса обучения) присутствуют в этом процессе «во множественном числе». С учётом сказанного, модель актуальна для проведения исследований ДСО, в частности, в связи с:

- разработкой методов синтеза, наполнения (адаптации), сопровождения и аттестации (контроля качества) ДСО;

- разработкой эффективных методов подачи и контроля предметного (прикладного, содержательного) материал в ДСО;

- сравнением и оценкой эффективности конкретных имеющихся ДСО;

- разработкой методов модифицирования ДСО в процессе их сопровождения, в направлении повышения их эффективности, и др.

2. Самоорганизующиеся карты Кохонена

Особенностью СКК является автономное обучение по схеме «без учителя» [3]. С точки зрения пользователя, это выглядит как процесс, в котором сама СКК «вырабатывает понимания» структуры данных. Результатом работы СКК является разделение данных на кластеры, которые могут далее непосредственно распознаваться пользователем. Соответственно, при поступлении входного образа, который ранее не предъявлялся сети, СКК относит его к одному из имеющихся кластеров, или (если режим обучения продолжается) определяет новый кластер.

Поскольку функциональным назначением («положительным эффектом» от работы) СКК является подготовка «системы кластеров» для конечного пользователя (человека, исследователя), то схема обучения «без учителя» не противоречит принципу диалоговости. Если СКК выдаёт результат, не воспринимаемый пользователем (сложный для интерпретации пользователем), то процедура обучения может быть повторена, либо модифицирована. Т.о., либо будет найдена и устранена ошибка в процедуре работы с СКК, либо пользователь изменит своё понимание результатов кластеризации (модифицирует свою интерпретацию). По существу, это и является диалогом.

Другим достоинством СКК является возможность визуализации многомерных данных – снижение размерности набора данных. При работе СКК, совокупность кластеров представляется в пространстве с меньшей размерностью. Если размерность является «антропоморфной» (соответствует практике обыденных человеческих представлений), результаты работы СКК могут использоваться человеком-оператором непосредственно, без дополнительного обобщающего анализа. Таким образом, СКК удобны в использовании в качестве интеллектуального че-

ловеко-машинного інтерфейса – усилителя определённых аспектов человеческого интеллекта.

В обоих описанных аспектах (кластеризация (группировка данных) и снижение размерности) СКК выступают как препроцессор входной информации, упрощающий работу человека-оператора и одновременно повышающий надёжность, не позволяющий оператору упустить из виду отдельные детали.

3. Параметры модели

ДПО – многопараметричен. Число параметров зависит от множества факторов. Вот некоторые из них:

- конкретная предметная область (изучаемый предмет);
- вид занятий (лекция, лабораторная работа, практические занятия и др.);
- манера изложения и уровень требований конкретного преподавателя;
- мотивированность, концентрация внимания;
- уровень подготовленности (образованности, способностей) аудитории;
- сложность изучаемого предмета;
- характер (степени структурированности) подаваемой информации.

Не все влияющие факторы могут быть одинаково легко выявлены (идентифицированы) применительно к конкретной реализации ДПО (конкретной предметной области). Соответственно, не все параметры ДПО могут быть известны (учтены). По этой причине, не все параметры могут быть адекватно реализованы в модели. Представленный далее вариант набора параметров может рассматриваться как частный ограниченный случай, соответствующий определённому начальному состоянию изученности вопроса и некоторому исходному пункту моделирования.

Предполагается наличие в модели трёх контуров, вложенных друг в друга (рис. 3). В каждом из контуров имеются логические (предметные) блоки V_{i1} , V_{i2} и блок принятия решений V_{i3} по выходу из вложенного контура и передаче управления внешнему контуру. Здесь i – номер контура; $i \in (1, 2, 3)$. Различие между блоками V_{i1} и V_{i2} – основная целевая (смысловая) обработка информации (V_{i1}) и вспомогательные функции (V_{i2}) по контурообразованию после принятия решения в V_{i3} . Конкретизация функций V_{i1} , V_{i2} и V_{i3} . – далее, с учётом прикладной специфики модели.

Внешний контур (V_{11} , V_{12} , V_{13}) – организация учебного процесса на уровне отдельных уроков обучения (циклов тренировки). Процесс обучения не может продолжаться непрерывно. Объективно имеется явление усталости мозга. Требуется промежуток для релаксации (перерывы, межурочные паузы), смысл которых – на биоинформационном уровне – рекреация отдельных нейронов мозга.

Объём и продолжительность временных периодов, задаваемых внешним контуром, устанавливается эмпирически и зависит от конкретных областей обучения (тренинга).

Соответствующие усреднённые значения (число уроков, продолжительность уроков и перерывов между ними и др.) хорошо известны из *человеческой* практики организации учебных процессов. Модель может (должна) их только воспроизводить, с переносом из реального времени в модельное.

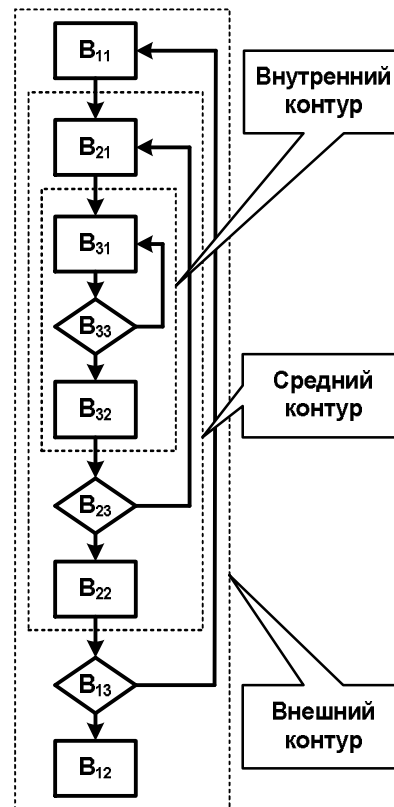


Рис. 3. Крупномасштабная структура модели

Как известно, реальное время в *человеческой* практике организации учебных процессов задаётся на основе физического времени, отмеряется по часам. При этом, в зависимости от сложности предмета, качества преподавания и уровня подготовленности аудитории, в «физическое время урока» вкладывается больший или меньший объём сообщаемых (закрепляемых) знаний. В модели – ситуация иная. Работу модели удобно организовывать исходя из фиксированности (предопределённости) сообщаемого информационного потока. Тогда неравномерность информационного потока, имитирующая «человеческий фактор», сама становится предметом моделирования. Более конкретно, в ключевых моментах, это рассмотрено далее.

Средний контур (V_{21} , V_{22} , V_{33}) – действие, происходящее внутри одного урока (сеанса обучения) – изображает последовательное предложение вопросов и получение ответов. Процесс предполагает

многократное повторение при предъявлении вопроса, что обусловлено спецификой фиксации информации в человеческом мозгу. В связи с этим, в модели задаётся фиксированное небольшое количество вопросов и фиксированное число предъявлений вопросов. Число предъявлений – приблизительно на порядок превышает число вопросов. Пример: ~20-30 вопросов и ~100-200 предъявлений вопросов – полученных ответов.

Внутренний контур (V_{31}, V_{32}, V_{33}) – организация (модельная имитация) процесса предъявления вопроса – получения и верификации ответа. Моделирование осуществляется с использованием *генераторов случайных чисел* (ГСЧ). Моделирование включает также учёт прежних значений, корректировку текущего состояния и изменение показателей обученности. Допустима, в частности, реализация исключения (изъятия) из цикла уже выученных вопросов.

В связи с реализацией представленной трёхконтурной структуры (рис. 3), для программной реализации модели вводится следующая система обозначений.

M – Массив вопросов, предъявленных обучаемому. Сами вопросы (в содержательном текстовом выражении) в модели не фигурируют и в массиве отсутствуют. Номера элементов массива – есть номера вопросов. Моделируются не конкретные содержательные ответы на конкретные вопросы, а изменение уровня изученности материала (массива вопросов) в процессе последовательного многократно повторяемого их предъявления. Изначальное (на момент запуска модели) заполнение массива M обозначает начальный уровень изученности материала. В частности, при полной новизне материала (отсутствии каких-либо изначальных знаний) – массив M изначально заполнен нулями.

В зависимости от того, какие проявления (особенности) процесса обучения исследуются конкретной версией (модификацией) модели, массив M может (должен) быть многомерным. Так, дополнительный слой массива M может изображать сложность вопроса – может быть заполнен значениями, изображающими *коэффициент сложности* (КС). При моделировании ситуации с вопросами одинаковой сложности, – значения КС одинаковые и слой КС – не нужен.

Если сложность вопросов различная, – ситуация может быть охарактеризована законом распределения значений КС. Тогда слой КС заполняется с использованием ГСЧ, воспроизводящего требуемый закон распределения. КС далее может учитываться при моделировании правильности ответов: вероятность правильного ответа понижается с ростом сложности вопроса.

N_M – число предъявлений вопросов в сеансе. Как отмечалось, им определяется продолжительность сеанса в виртуальном времени, т.е. продолжи-

тельность *сеанса обучения* (СО), которая, таким образом, является константой.

Допустим так же следующий вариант. Много-слойный массив, причём число измерений массива M превышает (включает в себя) N_M , т.е. массив хранит реакцию на предъявляемые вопросы, позволяет регистрировать всю динамику обучения.

N_S – есть число СО. Оно же – продолжительность курса. После ввода начальных данных (описаны выше), начинается процесс обучения. В нём – самый внешний цикл – по $i \in \{1, N_S\}$, внутри него – цикл по $j \in \{1, N_M\}$.

Внутренний цикл завершается внесением в протокол среза процесса обучения, а так же (либо в кратком варианте – только его) – текущего *коэффициента изученности* (КИ).

Коэффициент межсеансового забывания (КМЗ) – пропорционален межсеансовому интервалу. Может быть постоянным (хорошо составленное расписание, регулярные занятия, дисциплинированность), или переменным (плохо составленное расписание, или недисциплинированность).

Другой ГСЧ – предназначен для выбора случайного вопроса для предъявления.

4. Алгоритм. Пошаговое описание

Представим пошаговое описание алгоритма модели ДПО, крупномасштабно изображённого на рис. 3. Перед началом выполнения программы (перед запуском модели) в тексте программы прописываются все необходимые параметры. Как отмечалось выше (п. 2.х) языком реализации модели выбран Python? Сильной стороной которого является интерпретационность. Т.е., программа (а стало быть и модель) легко модифицируется, с заменой отдельных значений и целых блоков данных. В отлаженном варианте – это не приводит к осложнениям. Результаты оказываются сопоставимыми и воспроизводимыми. Дополнительное удобство состоит в простоте смены планов экспериментов, с учётом текущего изменения результатов.

Шаг 1. Пуск. Программа (модель) запускается на выполнение.

Шаг 2. Ввод начальных данные. Начальными (исходными) данными являются массив вопросов M о структурой слоёв согласно конкретной реализуемой версии модели; задаваемые наперёд число сеансов обучения и число повторов в каждом сеансе, а также коэффициенты (сложности, изученности, межсеансового забывания и др.), задаваемые в данной конкретной версии модели.

Шаг 3. Открытие протокола. Результаты моделирования должны протоколироваться (выводиться с виде протокола). Для этого при открытии протокола генерируется или указывается наперёд заданное его имя файла, режим протоколирования

(«write») – запись с –возможным уничтожением предыдущего содержимого файла, или «append») – дописывание в продолжение («в хвост») файла). Запись в протоколе снабжается датой и временем.

Шаг 4. Начало очередного сеанса курса обучения. Увеличение (приращение) на единицу значения счётчика числа сеансов курса обучения.

Шаг 5. Корректировка состояния обученности (содержимого массива M) с учётом фактора забывания (коэффициента межсеансового забывания). Предполагается, что между сеансами (отдельными уроками) имеются перерывы. За время перерыва – происходит восстановление свойств нейронов мозга, но вместе с тем частично забывается ранее усвоенная информация.

Шаг 6. Начало очередного цикла сеанса обучения (очередного урока). Увеличение (приращение) на единицу значения счётчика числа циклов (уроков) сеанса обучения.

Шаг 7. Тело сеанса обучения. Как описан выше (п. 3.2), сеанс обучения (урок) представляет собой последовательные многократные (с повторениями) предъявления единиц учебного материала. Настоящий Шаг представляет собой обработку одного предъявления. Единица материала предъявляется учащемуся – учащийся реализует контрольную операцию – система верифицирует правильность ответа и изменяет (повышает или понижает) значение переменной (элемента массива M), обозначающей уровень освоенности предъявленной единицы материала.

Шаг 8. Проверка завершения сеанса обучения. Условие завершения – реализация всех N_M предъявлений вопросов (единиц учебного материала) текущего сеанса обучения.

Шаг 9. Вывод в протокол результатов сеанса обучения: текущего состояния «уровня изученности» по каждой из единиц материала, а также усреднённых обобщающих значений для интегральной характеристики процесса обучения.

Шаг 10. Проверка завершения курса обучения. Условие завершения – реализация всех N_S сеансов обучения.

Шаг 11. Закрытие протокола. В текст протокола вносится запись о завершении протокола. Файл протокола закрывается.

Шаг 12. Стоп. Программа останавливается. Имеется протокол полученных результатов, которые могут обрабатываться в любом подходящем математическом пакете.

5. Обсуждение

Модель и полученные результаты – в целом вполне соответствуют интуитивным представлениям о прохождении процесса обучения. Полученные результаты демонстрируют рост изученности по мере проведения сеансов (уроков) обучения. При этом малое число предъявлений вопросов обеспечивает слабый рост изученности, большое число предъявлений – существенный прирост. Модель можно было бы видоизменить, так чтобы при достижении некоторого порогового значения вопрос снимался с рассмотрения. Более тонкий вариант – вопрос должен не сниматься с рассмотрения, а снижаться по частоте предъявления. В перспективе использования модели эти варианты обретут реализацию. Подобные изменения, сами по себе, являются параметрами модели и их влияние – есть отдельный предмет исследования.

Список литературы

1. Михаль, О.Ф. Глобально-исторический контекст развития средств вычислительной техники / О.Ф. Михаль // Бионика интеллекта : науч.-техн. журн. – Х.: ХНУРЭ, 2014. – Вып. 1 (82). – С. 55–62
2. Михаль О.Ф. Эволюционирование мультиагентной системы как аналог формирования индивидуального человеческого интеллекта / О.Ф. Михаль // Бионика интеллекта. – Х.: ХНУРЭ, 2016. – Вып. 2 (87). – С. 42–47.
3. Дяченко В.А. Интеллектуальный аспект обучения модифицированных самоорганизующихся карт Кохонена / В.А. Дяченко, О.Ф. Михаль // Бионика интеллекта: науч.-техн. журн. – Х.: ХНУРЭ, 2015. – Вып. 2 (85). – С. 35–40.

Надійшла до редколегії 9.02.2017

Рецензент: д-р техн. наук, проф. І.В. Рубан, Харківський національний університет радіоелектроніки, Харків.

МОДЕЛЮВАННЯ ПРОЦЕСУ ДІАЛОГОВЕ НАВЧАННЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ, ЩО САМООРГАНІЗУЮТЬСЯ

В.О. Дяченко, О.Ф. Міхаль

Авторами розроблено методологічні принципи з доведенням до конкретних структурних модельних схем: створення моделі, засобами якої можна відтворити динаміку процесу навчання ІСС, виходячи з його структури, безвідносно до конкретного змісту.

Ключові слова: модель, інтелектуальна система, що самоорганізується.

MODELING OF THE PROCESS OF DIALOGICAL TRAINING OF INTELLIGENT SELF-ORGANIZING SYSTEMS

V.O.. Dyachenko, O.F. Michal

The authors developed methodological principles with bringing to concrete structural model schemes: the creation of a model by means of which it is possible to reproduce the dynamics of the ASC training process, proceeding from its structure, without regard to the concrete content.

Keywords: model, intellectual self-organizing system.