На правах рукописи

КОРОБКИН АЛЕКСАНДР АЛЕКСАНДРОВИЧ

Разработка моделей и методов принятия решений с применением искусственного интеллекта для организации учебного процесса

05.13.17 – теоретические основы информатики

Автореферат диссертации

на соискание ученой степени
кандидата физико-математических наук

10 ДЕК 2009

ВОРОНЕЖ – 2009

Работа выполнена на кафедре информатики и методики преподавания математики Воронежского государственного педагогического университета.

Научный руководитель: доктор технических наук, профессор Астахова Ирина Федоровна

Официальные оппоненты:

доктор физико-математических наук, профессор Бугаев Юрий Владимирович

кандидат технических наук, доцент Курченкова Татьяна Викторовна

Ведущая организация: Воронежский государственный архитектурно-строительный университет

Защита состоится «23» декабря 2009 г. в 15 часов 10 минут на заседании диссертационного совета Д 212.038.20 при Воронежском государственном университете по адресу: 394006, г. Воронеж, ВГУ, математический факультет, Университетская пл., 1, ауд. 333.

С диссертацией можно ознакомиться в научной библиотеке Воронежского государственного университета.

Автореферат разослан «18» ноября 2009 г.

Ученый секретарь
диссертационного совета Д 212.038.20
кандидат физ.– мат. наук,
доцент Провоторов В.В.

Общая характеристика работы

Актуальность темы. Современные технические средства позволяют организовывать и планировать учебный процесс с использованием моделей, методов и алгоритмов искусственного интеллекта. В системе качества образования одним основных критериев выступает оптимальность учебных планов, расписаний и других составляющих учебного процесса. Для возможности применения моделей, методов и алгоритмов систем искусственного интеллекта необходимо формализовать учебный процесс.

Большинство исследований 70-80-х годов, направленных на решение задачи синтеза учебного плана, базировались на основе дерева подготовки специалистов и опирались на теорию графов (Гусев И.Т., Анисимов Б.В., Савельев А.Я., Горбатова Р.Е. и др.). Эти исследования оперировали низким количеством требований, налагаемых на учебный план, что было вызвано небольшой мощностью ЭВМ. Начиная с 90-х годов, появилось много разработок, основанных на принципе модульного обучения (Алексеева Ю.Н., Юсавичене П. и д.р.). Организация модульного обучения приводит к обособлению модулей и дублированию материала, что значительно сокращает объем учебного плана. В последние годы ведутся разработки, основанные на существовании связей между модулями. В работе Трофимовой О.К. рассматривается не только наличие связи между модулями, но и коэффициент тесноты этой связи. Несмотря на преимущества составления учебного плана из модулей на основе тесноты связей между ними, его синтез производится «с нуля», хотя в настоящее время система образования нуждается в средствах, позволяющих корректировать учебный план в соответствии с изменяющимися задачами обучения. Следовательно, возникает проблема синтеза рабочего учебного плана, основанного на использовании его типового аналога. Часть работы посвящена решению этой актуальной задачи.

Развитие исследований, направленных на решение задачи составления расписания можно разбить на два этапа. Первый этап имеет начало в 80-е годы и заканчивается в середине 90-х. В этот период масштабно применяются классические методы решения задач целочисленного программирования: метод полного перебора, метод раскраски графа, метод ветвей и границ (Безгинов А.Н., Трегубов С.Ю., Логоша Б.А., Петропавловская А.В., Гусева Н.Я. и д.р.). Используемые в этих разработках методы имеют крайне высокую степень формализации, как самой задачи, так и используемых алгоритмов. Применение классических методов в системах обучения становится малоэффективным, т.к. с увеличением размерности решаемой задачи существенно вырастают временные затраты, что не позволяет провести анализ влияния существующих факторов на решение. Это привело к появлению методов,

получивших название интеллектуальных. В их основе лежит использование различных эвристик и эвристических алгоритмов (Костин Л.А., Клеванский Н.Н., Маслов М.Г.). Решение задачи составления расписания с помощью эвристик не гарантирует нахождения глобального оптимума. Существует ряд работ, использующих для автоматизации составления расписания, математический аппарат нечеткой логики (Ханов Г.В., Алабужев Е.В., Борисов А.Н., Алексеев А.В., Меркурьев Е.В. и д.р.). Нечеткая логика позволяет заметно упростить формализацию требований, но часто приводит к построению расписания, имеющего не лучшие характеристики в результате перехода от «жестких» требований к более «мягким». В настоящее время для решения задачи составления расписания применяется ещё один новый подход – нейронные сети (Пилиньский М. Рутковская Д.). Важнейшим недостатком применения этого подхода является сложность выбора начального состояния нейронной сети. В последние годы особое распространение получили исследования методов эволюционного поиска (Ерунов В.П., Морковин И.И. Каширина И.Л., Низамова Г.Ф.). Применение методов эволюционного поиска приводит к получению хороших результатов, однако имеет место высокая вычислительная трудоёмкость и относительная неэффективность на заключительных этапах эволюции. В работе Низамовой Г.Ф. используются методы системного анализа, что позволяет упростить решаемую задачу, но это приводит к жесткой привязке составленного расписания к преподавательскому составу. На основании анализа и выявленных недостатков существующих разработок, в настоящей работе проводится исследование, направленное на решение задачи составления расписания с использованием агрегативного генетического алгоритма.

В последние годы развитие прогностики как науки привело к созданию множества моделей и методов прогнозирования, неравноценных по своей значимости. Применение интуитивных методов основано на получении экспертных оценок и обусловлено двумя противоположными ситуациями: либо объект слишком прост, либо сложен настолько, что аналитически не удается учесть влияние многих факторов (Рустыльник Е.И., Сысоев В.В., Чирко М.С., Самойлов М.С., и д.р.). В настоящее время большое распространение имеют разработки, основанные на использовании нейронных сетей (Алексеев В.И. Максивом А.В., и д.р.). Существенным недостатком использования нейросетевых методов является неопределенность при выборе начальных составляющих нейронной сети. Настоящая работа содержит разработку прогностической модели, основанной на методах нечеткой логике, что позволяет избежать требования больших объемов исходной информации, а также сложности их представления.

В настоящее время отсутствует единая теория формализации учебного процесса по многим компонентам: синтез рабочего учебного

плана, составление оптимального расписания, разработка прогностической модели развития учебного заведения. Исследование направлено на решение задачи формализации учебного процесса по нескольким составляющим с использованием алгоритмов и методов искусственного интеллекта.

Цель работы и основные задачи. Целью диссертационной работы является разработка и исследование моделей формализации учебного процесса с использованием методов и алгоритмов искусственного интеллекта. Для достижения цели необходимо решить следующие задачи:

1. Разработать модели учебного плана, и генетический алгоритм для получения рабочего учебного плана из его типового аналога.
2. Разработать модели расписания занятий и генетический алгоритм, осуществляющий поиск его оптимального варианта.
3. Построить прогностическую модель развития учебного заведения, основанную на методах нечеткой логики.
4. Разработать специальное программное обеспечение, реализующее разработанные методы.

Методы исследования. При решении поставленных задач использовались методы системного анализа, методы и алгоритмы эволюционного моделирования и нечеткой логики, методы объектно-ориентированного программирования.

Научная новизна. В работе получены следующие результаты, характеризующиеся научной новизной:

1. Модель учебного плана и генетический алгоритм, отличающийся возможностью синтезировать рабочий учебный план на основе его типового аналога.
2. Модель расписания и агрегативный генетический алгоритм, организующий поиск оптимального расписания, отличительными особенностями которого являются работа с агрегированными объектами и механизм направленной мутации.
3. Прогностическая модель развития учебного заведения, основанная на методах нечеткой логики, отличительной особенностью которой является функция ее настройка с помощью генетического алгоритма.
4. Специальное программное обеспечение, отличающееся объединением описанных методов в единую систему.

Теоретическая и практическая ценность. Работа имеет теоретический и практический характер. В работе проведена формализация учебного процесса: синтез учебного плана, составление

расписания и разработка прогностической модели. Практическая ценность работы заключается в возможности использования разработанного программного обеспечения для принятия решений при построении рабочего учебного плана, оптимального расписания учебных занятий и прогностической модели развития ВУЗа. Результаты работы используются в учебном процессе Лискинского филиала ВГУ. По результатам работы подана заявка на регистрацию программы на ЭВМ «Методист» в Федеральный институт промышленной собственности (ФИПС) от 13 ноября 2009 года.

Апробация работы. Основные результаты, представленные в диссертационной работе, докладывались и обсуждались на Международной конференции «Современные проблемы механики и прикладной математики» (Воронеж, 2007), на III Международной научной конференции «Современные проблемы математики и математического моделирования» (Воронеж, 2009), на Математической конференции молодых преподавателей и студентов Лискинского филиала ВГУ (Воронеж, 2007), на научных сессиях Воронежского государственного университета, на научных сессиях Воронежского государственного педагогического университета (Воронеж, 2009).

Публикации. Результаты диссертации опубликованы в 8 работах. Из совместных работ в диссертацию вошли только результаты, принадлежащие лично диссертанту. Списку ВАК соответствует работа [8].

Структура и объём диссертации. Диссертация состоит из введения, 5 глав, разбитых на пункты, заключения, списка используемой литературы из 130 наименований, и приложения. Общий объем диссертации – 144 страницы. Работа содержит 39 рисунков, 3 диаграммы и 1 таблицу.

Краткое содержание работы

Первая глава содержит в себе обзор существующих методов и разработок, направленных на решение поставленных задач. Производится сравнительный анализ существующих разработок, направленных на решение задачи синтеза учебного плана, выявлены наиболее значимые недостатки и противоречия в них. В работе проводится обзор существующих методов решения задачи составления расписания, и приводятся положительные аспекты применения моделей и методов искусственного интеллекта в области принятия решений при составлении расписания. Работа содержит общую классификацию существующих методов прогнозирования, проводится анализ классических методов

прогнозирования, выявляются слабые стороны и трудности их применения. Определенное место в исследовании занимает анализ существующих интеллектуальных моделей и методов, применимых к построению модели прогнозирования.

Во второй главе описана одна из составляющих учебного процесса – формализация составления рабочего учебного плана, предлагается математическая модель учебного плана, содержащая ограничения, налагаемые на него, проведена структуризация исходной информации. Разработан и описан генетический алгоритм, позволяющий находить оптимальный состав и структуру рабочего учебного плана. Приведены результаты работы описанного алгоритма.

Математическая постановка начинается с определения возможного содержания обучения, представленного в виде множества дисциплин, которые в свою очередь состоят из учебных блоков.

$$MOD = \{a_{i,j}\}, i = \overline{1,N}; j = \overline{1,m(i)},$$

где $a(i,j)$ – $j-ый$ модуль $i-ой$ дисциплины;

N – количество дисциплин;

$m(i)$ – количество модулей в $i-ой$ дисциплине.

Учебным рабочим планом (УРП) на d дискретных интервалах времени называется множество модулей $a(i,j)$:

$$УРП = \{a(i,j) \in MOD | V_{УРП} \le V_{MOD}\},$$

где V_{MOD} – объем всех модулей множества MOD; $V_{УРП}$ – объем модулей, принадлежащих подмножеству рабочего учебного плана.

Тогда учебным планом будет являться некоторое подмножество множества всех дисциплин, суммарный объём блоков которого, не превышает допустимый объём учебного плана.

Модуль $a(i,j)$ представляется следующим образом:

$a(i,j) = \{a(i,j,1), a(i,j,2), a(i,j,3), a(i,j,4), a(i,j,5), a(i,j,6)\}$,где $a(i,j,1)$ – объем лекционных занятий, $a(i,j,2)$ – объем практических занятий, $a(i,j,3)$ – объем лабораторных занятий, $a(i,j,4)$ – коэффициент обобщенной значимости модуля, $a(i,j,5)$ – начало изучения модуля, $a(i,j,6)$ – конец изучения модуля,

Для учета взаимосвязей между модулями вместе с коэффициентом тесноты этой связи предусматривается наличие квадратной матрицы М, у которой номером строки является номер элемента модуля дисциплины-родителя, а номера столбцов представлены номерами модулей дисциплин-потомков. Элементами матрицы являются коэффициенты тесноты связи между «родственными» дисциплинами. Коэффициент тесноты связи определяется из отрезка [0,1] экспертами и определяет, насколько существенен порядок следования для двух рассматриваемых модулей.

Ограничения на учебный план выглядят следующим образом:

1. Календарное время окончания любого модуля не должно превышать установленного срока d обучения в вузе.

$$\max_{a(i,j)\in УПЛ}\{a(i,j,6)\} \le d$$

2. Количество дисциплин N в плане не более K.

$$N \le K, \quad L = \varepsilon\{i \mid a(i,j,5) \ge 0 \wedge a(i,j,6) \le d\}, \quad N = |L|$$

3. Количество дисциплин в любом семестре не более KS.

$$\forall s[L(s) \le KS], \quad L(s) = \varepsilon\{i \mid a(i,j,5) \ge n(s) \wedge a(i,j,6) \le k(s)\}, \quad LS = |L(s)|$$

4. Начало и окончание изучения любого учебного модуля должно находиться "внутри" какого-либо семестра.

$$\forall i \forall / \exists s[n(s) \le a(i,j,5) \wedge k(s) \ge a(i,j,6)]$$

Целевая функция представляет собой функцию минимизации суммарного значения штрафных показателей и нарушения ограничений, помноженных на их весовые коэффициенты. Штрафными показателями являются: нарушение объемов занятий в течение одного семестра, нарушение равномерности распределения количества дисциплин, нарушение порядка следования модулей дисциплин с учетом тесноты связей между ними.

$$F_ц = \underset{i=1..количество\ штрафов}{\sum}(sh_i \cdot k_sh_i) + \underset{j=1..количество\ ограничений}{\sum}(og_j \cdot k_og_j)$$

Задача составления учебного плана представляет собой задачу выбора из всего объёма дисциплин наиболее важных для данной специальности и расположения их по семестрам и неделям, оптимальным для выбранного критерия образом.

Исходными данными в решении задачи составления рабочего учебного плана являются дисциплины обучения, таблица тесноты связей, таблица базовых дисциплин и временные нормы. Данную информацию удобно представить в виде реляционной базы данных. Это позволит организовать довольно простую выборку необходимых данных при решении задачи, а также иметь не сложные способы редактирования и занесения новой информации в базу данных учебного плана. Проектирование описанной базы данных осуществлено с помощью ER – диаграмм.

На основе анализа предметной области для представления решения была выбрана эволюционная модель, суть которой состоит в следующем: популяция решений будет состоять из определённого количества особей, каждая из которых обладает одной хромосомой рис. 1. Популяция представляется в виде двумерного массива. Хромосома состоит из генов, которые определяют порядковый номер семестра. Массив, хранящий информацию о блоках дисциплин, представляет собой набор всех необходимых для изучения блоков дисциплин. Вспомогательный массив организует взаимно – однозначную связь между содержанием гена и

номером блока дисциплины. Количество генов определяется количеством учебных блоков.

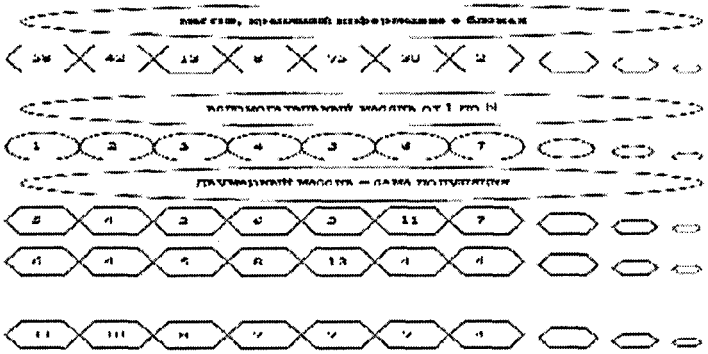

Рис. 1 Схема популяции.

Генетический алгоритм состоит из следующих шагов:

1. формирование начальной популяции;
2. селекция особей;
3. скрещивание особей с большими значениями функции пригодности;
4. операция мутации над потомством;
5. отбор особей в новую популяцию;
6. проверка критерия остановки алгоритма;
7. выбор наилучшей особи.

В рамках данного алгоритма применяется многоточечный оператор мутации, причем количество точек мутации также случайно, как и позиции разыгрываемых генов. Благодаря такому подходу, состав популяции становится разнообразнее, что позволяет расширить пространство поиска оптимального решения.

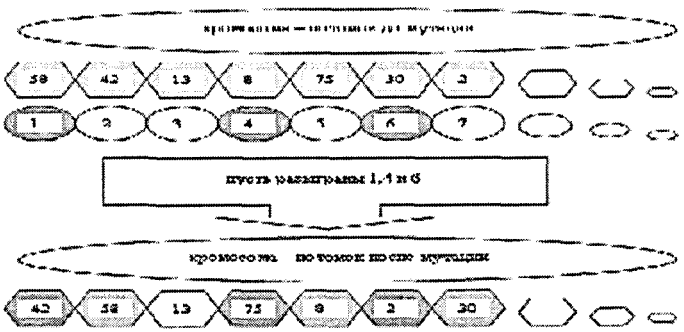

Рис. 2 Многоточечный оператор мутации

Одна итерация генетического алгоритма выглядит в виде пяти связанных этапов. Во время каждой итерации численность популяции

увеличивается вдвое, а потом возвращается в исходное состояние, т.е. поиск оптимального решения производится в более широком пространстве, нежели размеры самой популяции. Это во многом определяется выбором генетических операторов. Итерационный процесс генетического алгоритма продолжается до тех пор, пока не пройдет заданное число поколений или не выполнится какой-либо иной критерий останова. Описываемый генетический алгоритм использует следующий критерий остановки. Алгоритм завершает свою работу, когда разница между значениями целевой функции «худшей» особи и «лучшей» практически не существует. Выбор именно этого критерия остановки алгоритма связан с тем, что при столь малой разнице значений целевой функции разных хромосом, становится очевидным, что дальнейшая работа алгоритма не принесет каких-либо ощутимых изменений в популяцию.

Получены диаграммы, отражающие динамику изменения численности популяции во время каждого этапа эволюционного поиска, диаграмма сходимости функции пригодности генетического алгоритма.

Третья глава содержит формализацию другой составляющей учебного процесса: разработку математической модели расписания, содержащей ограничения и требования к нему. В этой главе содержится описание генетического алгоритма, основанного на использовании агрегированных объектов и механизме направленной мутации.

Теоретико-множественная модель расписания представляется в виде функции R, отображающей декартово произведение множеств на множество нулей и единиц {0,1}:

$$R: G \times A \times D \times P \times T \to \{0,1\}.$$

Если в указанной группе проводятся G занятия в аудитории A по дисциплине D, преподавателем P, во время учебной пары T , то функция принимает значение равное 1, в противном случае – 0.

В работе используется системный подход, в основе которого используется рассмотрение объекта как целостного множества элементов в совокупности отношений и связей между ними, т.е. рассмотрение объекта как системы. Основными принципами системного подхода являются: целостность; иерархичность строения; структуризация; множественность. На основе системного подхода в работе производится декомпозиция исходного объекта на два менее сложных, решение общей задачи сводится к поиску решения более мелких задач на каждом этапе синтеза.

Для дальнейших рассуждений, приводящих к построению математической модели, проводится детальный анализ взаимодействующих объектов предметной области.

Для описания аудитории, расположенной в $i-ом$ корпусе, под номером j и имеющей тип t вводится в рассмотрение трехкомпонентный кортеж:

$$A = \{a_j\},\ a_j = \left(a_j^v, a_j^s, a_j^{type}\right) \quad (1.1)$$

Временные интервалы описываются множеством T, каждый элемент которого представляет собой трехкомпонентный кортеж вида:

$$t_k = \left(t_k^w, t_k^d, t_k^p\right) \quad (1.2)$$

Объект «преподаватели» представляется в виде множества векторов следующей структуры:

$$P = \{p_i\},\ p_i = \{d_1, d_2, ..., d_m\}, \quad (1.3)$$

где p_i – код преподавателя, d_j – одна из дисциплины, читаемых этим преподавателем, а m – общее количество предметов, которые он может читать.

Множественное описание объекта «занятие» представляется следующим образом:

$$Z = \{z_i\},\ \{z_i\} = \left\{z_i^d, z_i^g, z_i^e, z_i^h, z_i^l, z_i^a\right\}, \quad (1.4)$$

где z_i – элемент множества блоком занятий $Z(i = \overline{1, N_{блоков}})$; $N_{блоков}$ – количество блоков занятий; z_i^d – дисциплина, преподаваемая в блоке; z_i^g – группа, у которой проводится занятие; $z_i^e \in \left\{\frac{1}{2}, 1\right\}$ – признак подгруппы; z_i^h – параметр, определяющий длительность блока; z_i^l – параметр определения вида занятия; z_i^a – индекс допустимого подмножества аудиторий.

В классических моделях принято использовании всех пяти множеств: множество аудиторий, множество учебных групп, множество преподавателей, множество дисциплин и множество временных промежутков времени – учебных пар, что приводит к сложности реализации поисковых алгоритмов.

На основе (1.1) – (1.3), математическая модель расписания учебных занятий представляется в виде трех векторов:

$$\alpha = \left(\alpha_1, \alpha_2, ..., \alpha_i, ..., \alpha_{N_{блоков}}\right).\ t = \left(t_1, t_2, ..., t_i, ..., t_{N_{блоков}}\right),\ \rho = \left(\rho_1, \rho_2, ..., \rho_i, ..., \rho_{N_{блоков}}\right), \quad (1.5)$$

где $\alpha_i \in A$ - код аудитории, назначенный блоку занятий $z_i \in Z$, $t_i \in T$ - код учебной пары, назначенный первому занятию из блока занятий $z_i \in Z$, $\rho_i \in P$ - код преподавателя, назначенный блоку занятий $z_i \in Z$.

Математическая модель расписания включает ограничения, налагаемые на него, а также ряд требований. Ограничения, налагаемые на расписание, описываются следующим образом:

1. Математическое выражение, описывающее отсутствие накладок для аудиторий выглядит следующим образом. Для каждой упорядоченной двойки элементов: аудитория и учебная пара существует либо единственный блок занятий из множества Z, что означает проведение занятия этого блока в этой аудитории в момент данной пары, либо отсутствие блока занятия, указывающее на то, что аудитория свободна.

$$\forall(a_i, t_j): a_i \in A, t_j \in T(\exists z_k : (a_i = a_k) \wedge (z_k \in Z^{t_j})) \vee (\neg\exists z_k : (a_i = a_k) \wedge (z_k \in Z^{t_j})), \quad (1.6)$$

где Z^{t_k} – множество блоков занятий, проводимых во время пары t_k.

2. Отсутствие накладок для преподавателей описывает, что для каждой упорядоченной двойки элементов: пара и преподаватель, существует либо единственный блок занятий, которые ведет данный преподаватель во время заданной пары, либо этого блока не существует вообще. Данное ограничение формализуется следующим образом:

$$\forall(p_i,t_j): p_i \in P, t_j \in T(\exists z_k : (p_i = p_k) \wedge (z_k \in Z^{t_j})) \vee (\neg\exists z_k : (p_i = p_k) \wedge (z_k \in Z^{t_j})), \quad (1.7)$$

где Z^{t_j} – множество блоков занятий, проводимых во время пары t_k.

3. Отсутствие накладок для учебных групп означает, что для каждой пары элементов: группа и пара, сумма компонент z_i^e вектора z_i блоков из множества $Z^{g_j} \bigcap Z^{t_k}$ не превышает единицы. Во время конкретной пары группа находится на одном занятии, или проводится занятие только у одной из подгрупп, либо у обеих, либо занятий нет вообще. Тогда математическое описание данного ограничения имеет вид:

$$\forall(g_n,t_j): g_n \in G, t_k \in T \sum z_i^e \le 1, i \in Z^{g_n} \bigcap Z^{t_j}, \quad (1.8)$$

где Z^{g_n} – множество блоков занятий, в которых присутствует группа g_n, а Z^{t_j} – множество блоков занятий, проводящихся во время пары t_k.

4. Соответствие типа аудитории проводимому занятию предполагает, что для каждого блока занятия $z_i, z_i \in Z$ аудитория выбирается из допустимого подмножества аудиторий, код этого подмножества хранит компонента z_i^a.

$$\forall z_i \in Z\, a_i \in A^{z_i^a}. \quad (1.9)$$

5. Ограничение, налагаемое на количество учебных пар, проводимых в течение одного учебного дня, означает, что для каждой пары элементов: группа и день, число проводимых пар не превышает максимально допустимого – $N_{пар_max}$.

$$\forall(b_\tau,g_n): b_\tau \in B, g_n \in G \sum z_i^e \le N_{пар_max}, i \in I_{g_n}^{b_\tau}, \quad (1.10)$$

где $B = \{b_1, b_2, ..., b_{N_{дней}}\}$ – множество учебных дней. Каждый элемент описанного множества, определяется следующим образом:

$$b_\tau = \{t_j \in T : t_j^d = b_\tau\}.$$

6. Требование отсутствия окон для учебных групп означает, что для каждой пары элементов: день и группа, количество пар, проводимых в этой группе в текущий день должно равняться величине:

$t_\max_\ number_{g_n}^{b_\tau} - t_\min_\ number_{g_n}^{b_\tau} + 1$, а также отсутствие окон у групп.

$t_\max_number_{g_n}^{b_\tau}$ – максимальный номер пары в течение дня b_τ для группы g_n, $t_\min_number_{g_n}^{b_\tau}$ – минимальный номер пары в течение дня b_τ, для группы g_n, $I_{g_n}^{b_\tau} = \{i : (z_i \in Z^{g_n}) \wedge (t_j^d = b\tau)\}$ – множество номеров блоков занятий, проводимых для группы g_n во время дня b_τ.

С учетом описанных выше формул, рассматриваемое ограничение выражается следующим образом:

$$\forall(b_t, g_n): b_t \in B, g_n \in G$$

$$(\sum_{i \in l^{b_t}_{g_n}} z^t_i = t_\max_number^{b_t}_{g_n} - t_\min_number^{b_t}_{g_n} + 1) \qquad (1.11)$$

$$\wedge(\forall t: t_\min l^{b_t}_{g_n} \le t \le t_\max l^{b_t}_{g_n} \sum_{i\,z_i \in Z^{g_n} \wedge l^p_i = t} z^t_i = 1)$$

С учетом описанной математической модели (1.5), а также при соблюдении ограничений (1.6) – (1.11), требуется найти такой вариант выбора векторов α, t, ρ, удовлетворяющее ограничениям (1.6) – (1.11), а также минимизирующее значение критерия потери качества K. Критерий качества основывается на желательных требованиях.

$$K = \varphi(\alpha, t, \rho) = \sum_{i=1}^{N} c_i w_i(\alpha, t, \rho),$$

где w_i – значение штрафного коэффициента за невыполнение $i-го$ требования, а w_i – оценка степени невыполнения $i-го$ желательного требования, а N – количество требований.

На основании описанных требований строится целевая функция на основе минимизации штрафных показателей. Каждое нарушение ограничения или желательного требование увеличивает значение целевой функции в соответствии с коэффициентом значимости требования. В результате целевая функция в общем виде описывается формулой:

$$F_ц = \underbrace{\sum_{i=1..N}}_{\text{количество ограничений}} og_i \cdot k_og_i) + K$$

Разработанный алгоритм состоит из следующих шагов:

- формирование начальной популяции;
- селекция особей;
- кроссинговер;
- операция случайных мутаций;
- операции направленных мутаций;
- отбор особей;
- возврат ко второму шагу, если не достигнуто решение;
- выбор наилучшей особи.

На рис. 3 изображена схема наиболее важного в данном алгоритме оператора направленной мутации.

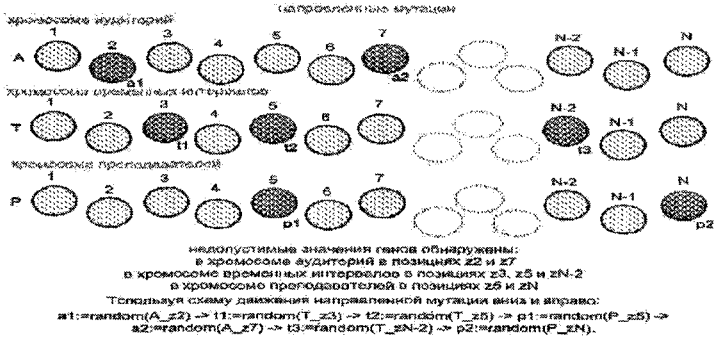

Рис. 3 Схема работы оператора направленной мутации.

Разработанный алгоритм предусматривает два критерия остановки:

- остановка при стабилизации значения функции пригодности;
- остановка алгоритма при незначительном изменении «лучшего» значения функции пригодности.

В четвертой главе содержится описание построения прогностической модели на основе нечеткой логики. Для настройки работы прогностической модели используется генетический алгоритм

В главе рассматривается весь процесс нечеткого управления, состоящий из последовательных действий: фаззификация, разработка нечетких правил и дефаззификация. Приведены описания и виды наиболее используемых функций принадлежности, рассмотрены понятия лингвистической и нечеткой переменных, а также понятие термов Нечеткий логический вывод производится с помощью алгоритма Мамдани.

Для разработки методов прогнозирования производится структуризация исходных данных предметной области, проводится анализ экспериментальных данных, необходимый для формирования термов лингвистической выходной переменной, выявляются закономерности прогнозируемой структуры, на основании которых происходит выработка высказываний на естественном языке, которые в дальнейшем формируются в правила базы знаний. На основе отобранных правил формируются функциональные связи, необходимые для построения сети прогнозирования.

В решаемой задаче используется аналитическая модель функции принадлежности переменной x произвольному нечеткому терму T в виде:

$$\mu^T(x) = 1/(1+((x-b)/c)^2), \text{ где}$$

b и c – параметры настройки. Параметр b представляет собой координату максимума функции принадлежности – $\mu^T(b) = 1$. Параметр c

является коэффициентом концентрации – растяжения функции. Число b представляет собой наиболее возможное значение переменной x для нечеткого терма T.

На основе экспертных оценок выходной лингвистической переменной формируются параметры функций принадлежности каждого терма.

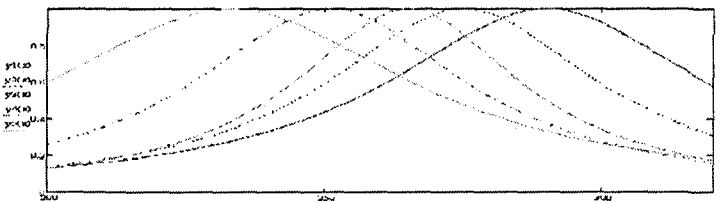

Рис. 4 Функции принадлежности лингвистических оценок до настройки.

На основания метода центра тяжести определяется процедура дефаззификации, направленная на получение четких значений выходной переменной. В случае дискретных величин универсального множества, для деффазификации нечеткого множества используется следующая формула:

$$X = \frac{\sum_{i=1}^{K} k_i \cdot \mu_A(X_i)}{\sum_{i=1}^{K} \mu_A(X_i)}.$$

Для повышения точности прогноза проведена процедура настройки модели прогнозирования с помощью генетического алгоритма. Суть его работы заключается в выборе значений функций принадлежности лингвистических оценок таким образом, чтобы отклонение прогноза от экспериментальных данных было минимальным.

Графики функций принадлежности лингвистических оценок приведены на следующем рисунке.

Рис. 5 Функции принадлежности лингвистических настроек после настройки

Результаты работы разработанной модели прогнозирования показаны на следующем рисунке.

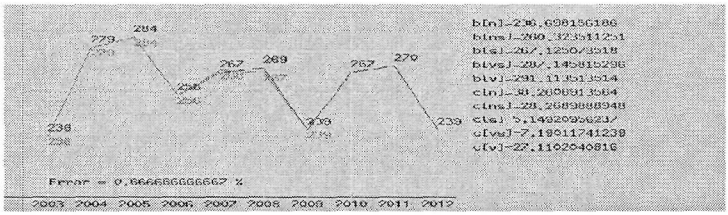

Рис. 6 Результаты прогнозирования после настройки модели

В пятой главе описана структура всего программного продукта. Описано взаимодействие основных функциональных блоков и обосновано использование связки PHP+MySQL+HTML. Отражена функциональная схема разработанного программного обеспечения. На основе функциональной схемы разработана структурная схема взаимодействия основных модулей.

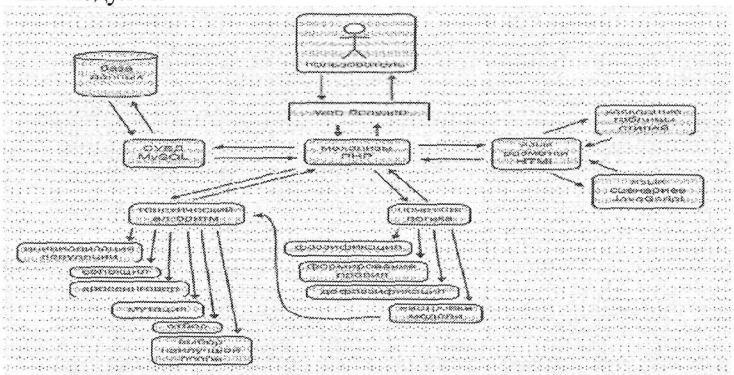

Рис. 7 Структурная схема взаимодействия основных модулей

Структура основного модуля работы с базой данных с применением языка PHP и СУБД MySQL имеет следующий вид:

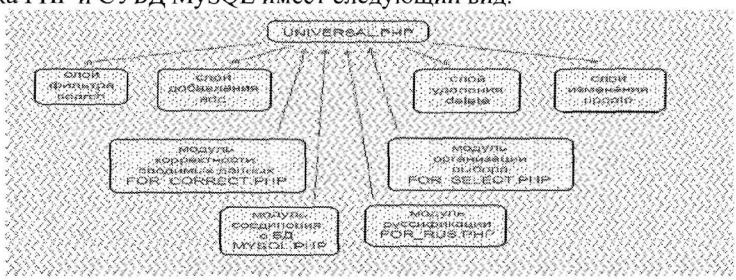

Рис. 8 Базовый модуль для работы с БД

Рассматриваются инструменты, организующие разработанные интеллектуальные методы принятия решений. Это инструмент синтеза учебного плана, находящийся в файле UCH_PLAN_1.PHP. Синтез расписания производится в результате исполнения сценария

RASPISANIE.PHP, изложены и отражены принципы функционирования модуля прогнозирования, основанного на методах нечеткой логики. Реализация этого метода описана в коде сценария PROGNOZ.PHP.

Заключение отражает основные результаты, полученные в ходе выполнения диссертационной работы.

Основные результаты работы

Основным результатом настоящей диссертационной работы являются: формализация учебного процесса, разработанные математические и эволюционные модели для составления рабочего учебного плана, расписания занятий, прогностическая модель развития учебного заведения.

Среди наиболее важных результатов можно выделить:

1. Проведенный анализ существующих методов решения задач составления учебного плана, задач составления расписания и методов прогнозирования.
2. Разработка математической модели учебного плана, а также реализация на ее основе генетического алгоритма составления рабочего учебного плана.
3. Разработка модели расписания на основе системного подхода и генетического алгоритма с механизмом направленной мутации. Применение агрегированных объектов оказалось предпочтительнее для поиска оптимального решения.
4. Разработка модели прогнозирования численности студентов, основанной на методах нечеткой логики, алгоритме нечеткого вывода Мамдани, и осуществление настройки модели с помощью генетического алгоритма.
5. Разработка специального программного обеспечения, обеспечивающего работу интеллектуальных методов принятия решений для поставленных задач.

Публикации автора по теме диссертации

1. Коробкин А.А. Автоматизированное рабочее место библиотекаря филиала/А.А.Коробкин, Е.Н.Ветров// Мат. Конф. Молодых преп. и студентов Лискинского филиала ВГУ. – Воронеж: Воронеж. госун-т, 2007. – 43-47.
2. Коробкин А.А. Интерфейс электронной web-ориентированной библиотеки филиала с конвертированием данных из формата UNIMARK/ А.А.Коробкин, И.Ф.Астахова// Современные проблемы

механики и прикладной математики труды. Межд. Конф.– Воронеж: Воронеж. госун-т, 2007. – Вып. 3.- С. 19-22.
3. Коробкин А.А. Электронная Web-ориентированная библиотека с конвертированием данных из формата UNIMARK/ А.А.Коробкин, Е.Н.Ветров// Труды молодых ученых. – 2007. Вып. 1-2.- С.16-24.
4. Коробкин А.А. Разработка моделей информационной системы построения расписания/ Коробкин А.А., Астахова И.Ф.// Современные проблемы математики и математического моделирования: Материалы III Международной научной конференции. Часть 2. – Воронеж:Научная книга,2009. - с. 156-158.
5. Коробкин А.А. Использование агрегативного генетического алгоритма для составления расписания учебных занятий ВУЗа. // Современные проблемы науки и образования - 2009.-№6. (приложение "Технические науки"). - с. 10.
6. Коробкин А.А. Применение генетических алгоритмов для проектирования учебных планов ВУЗов/Коробкин А.А. //Новые технологии в образовании №4,В:Мастеринг ВГПУ,2009. с. 69-71.
7. Коробкин А.А. Использование агрегативного генетического алгоритма для сотавления расписания учебных занятий ВУЗа/Коробкин А.А.//Перспективы развития современной школы №4,Воронеж:Мастеринг ВГПУ,2009. с. 32-35.
8. Коробкин А.А. Использование агрегативного генетического алгоритма для составления расписания/ Коробкин А.А. //Вестник Воронежского государственного технического университета, том 5 ,№11,2009. с. 184-186.

Подписано в печать 18.11.09. Формат 60×84 $^{1}/_{16}$. Усл. печ. л. 1.2
Тираж 110 экз. Заказ 1899

Отпечатано с готового оригинала-макета
в типографии Издательско-полиграфического центра
Воронежского государственного университета.
394000, Воронеж, ул. Пушкинская, 3.