

На правах рукописи

Молодченков Алексей Игоревич

**МОДЕЛИ, МЕТОДЫ И АРХИТЕКТУРЫ
ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ
ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ (НА
ПРИМЕРЕ МЕДИЦИНЫ И ПСИХОЛОГИИ)**

Специальность: 05.13.01 – Системный анализ, управление и обработка
информации (информационно-вычислительное обеспечение)

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук**

Москва, 2017

Работа выполнена в Федеральном государственном учреждении «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН), лаборатория №0-2 «Динамические интеллектуальные системы».

Научный руководитель: доктор физико-математических наук, профессор Осипов Геннадий Семенович

Официальные оппоненты: Грибова Валерия Викторовна
доктор технических наук, Федеральное государственное бюджетное учреждение науки Институт автоматики и процессов управления Российской академии наук
Дальневосточное управление, Заместитель директора по научной работе

Шаповалов Валентин Викторович
доктор технических наук, профессор, Федеральное государственное автономное образовательное учреждение высшего образования Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), директор
Федеральное государственное унитарное предприятие «Российский научно-технический центр информации по стандартизации, метрологии и оценке соответствия» (г. Москва).

Защита состоится «__» _____ 2017 г. в 11:00 на заседании диссертационного совета Д 002.073.04 при Федеральном государственном учреждении “Федеральный исследовательский центр “Информатика и управление” Российской академии наук” (ФИЦ ИУ РАН) по адресу: 117312, Москва, проспект 60-летия Октября, 9 (конференц-зал, 1-й этаж).

С диссертацией можно ознакомиться в библиотеке ФИЦ ИУ РАН по адресу: Москва, ул. Вавилова, д. 40.

Электронные версии диссертации и автореферата размещены на официальном сайте ФИЦ ИУ РАН по адресу: <http://www.frccsc.ru>

Электронная версия автореферата размещена на официальном сайте ВАК Министерства образования и науки РФ по адресу: <http://vak.ed.gov.ru>

Отзывы и замечания по автореферату в двух экземплярах, заверенные печатью, просьба высылать по адресу: 117312, Москва, проспект 60-летия Октября, 9, ФИЦ ИУ РАН, диссертационный совет Д 002.073.04

Автореферат разослан «__» _____ 2017 г.
Телефон для справок: +7(499) 135-51-64

Ученый секретарь диссертационного совета
доктор технических наук, профессор

В.Н.Крутько

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Процессный подход к организации управления предприятиями в различных отраслях экономики сегодня признан наиболее перспективным. В последние годы появилось немало работ по изучению возможностей использования систем управления бизнес-процессами для автоматизации медицинских технологических процессов (МТП). Возник и соответствующий термин – «*careflow*» по аналогии с «*workflow*», означающий поток работ, направленных на оказание медицинской помощи.

Сегодня существует три основных направления исследований в области автоматизации медицинских технологических процессов:

- 1) электронные медицинские карты в различных вариантах (включая карты здоровья, карты пациентов);
- 2) поддержка принятия решений на основе клинических руководств;
- 3) управление клиническими процессами (электронные клинические пути – э-КП).

Электронные медицинские карты (ЭМК) постепенно преобразуются из простого хранилища клинических данных о пациенте в многофункциональную систему, позволяющую организовывать, распределять и использовать медицинские знания с высокой степенью защищенности, облегчать и регулировать взаимодействие всех участников лечебно-диагностических процессов

Исследования в области систем поддержки принятия решений (СППР) в последние два десятилетия были направлены, в основном, на разработку компьютеризированных клинических руководств. СППР на основе клинических руководств призваны представлять клиницистам контекстные рекомендации в нужный момент, обеспечивать взаимосвязь с ЭМК и другими информационными системами, осуществлять мониторинг этапов оказания медицинской помощи

Третье направление исследований берет начало от концепции клинических путей (*clinical pathways*), предложенной в 80-х годах XX века в качестве метода управления клиническими случаями и медсестринской помощью в Медицинском центре Новой Англии, в Бостоне. Клинический путь представляет собой пациент-ориентированный план лечебных мероприятий, которые должны быть выполнены за один эпизод оказания медицинской помощи. КП имеют мультидисциплинарный характер и применяются для внедрения клинических руководств в клиническую практику, а также для снижения нежелательных отклонений от медицинских стандартов.

Несколько иным образом дело обстоит в психологии. Основные усилия по внедрению информационных технологий направлены на разработку компьютеризированных психодиагностических методик, экспертных систем, интерпретаторов результатов тестирований, моделей прогноза, психологического сопровождения конкретных видов деятельности.

Анализ существующих систем и методов моделирования и поддержки технологических процессов в области медицины и психологии показал, что они ориентированы на поддержку лечения заболеваний, в не конкретного пациента (клиента в психологии). Более того во всех системах, направленных на поддержку технологических процессов в области медицины и психологии отсутствуют средства оценки качества оказанной медицинской и психологической помощи, анализа отклонений от технологических процессов, выявление причин отклонений от процесса лечения, средств управления технологическими процессами и адаптации МТП под конкретного пациента. Настоящая работа посвящена разработке методов, алгоритмов и архитектур программных средств поддержки технологических процессов в области медицины и психологии, что свидетельствует о её актуальности.

Цели и задачи исследования

Целью диссертационной работы является разработка новых методов и алгоритмов анализа и поддержки технологических процессов в области медицины и психологии.

Для достижения поставленной цели в работе решены следующие задачи:

- выполнен анализ существующих методов и программных средств поддержки технологических процессов в области медицины и психологии;
- предложены способы представления технологических знаний с использованием онтологий и неоднородных семантических сетей;
- разработан новый алгоритм автоматического синтеза обобщенной схемы технологических процессов на основе прецедентной информации;
- предложены архитектуры программных средств поддержки технологических процессов в области медицины и психологии.

Методы исследования. В диссертационной работе использованы методы искусственного интеллекта, машинного обучения, теории множеств, математической логики, интеллектуального анализа данных, интеллектуального анализа процессов.

Научная новизна.

Современные подходы в области поддержки медицинских технологических процессов направлены на разработку систем, решающих узкий круг задач. К таким системам относятся системы поддержки принятия решений, формализации клинических руководств, лекарственных назначений и др. Аналогичная ситуация складывается и в практической психологии, где все разработки направлены на компьютеризацию диагностических методик в психодиагностике и др. В последние годы наблюдаются попытки моделирования медицинских технологических процессов. Однако построенные модели направлены на лечение заболевания, а не больного. В отличие от этого, предлагаемые в настоящей работе методы позволяют синтезировать технологические процессы, являющиеся персонализированными и направленными на лечение больного, а не заболевания.

Такой подход становится возможным благодаря тому, что широкое применение медицинских информационных систем позволяет накапливать

информацию о лечении нозологических форм конкретных пациентов. Эта информация содержит данные о состояниях больных, лечебно-диагностических мероприятиях, проводимых с больными, и др. Если эта информация становится доступной, то можно строить обобщенные схемы медицинских технологических процессов, которые будут учитывать индивидуальные особенности пациентов, а на их основе персонализированные МТП.

Исследователи в области анализа процессов занимаются разработкой методов и алгоритмов автоматического построения моделей технологических процессов на основе прецедентной информации. Однако современные алгоритмы автоматического построения моделей технологических процессов только частично решают проблемы петель, коротких циклов, несвободного выбора, устранения шумов и др., что отражается на качестве их результатов. Решение этой задачи также является одной из целей настоящей работы.

Основные результаты работы:

- впервые предложено матричное представление медицинских технологических процессов;
- доказаны утверждения об ассоциативности и не коммутативности последовательного применения произвольного числа операторов МТП;
- введена операция покомпонентного сложения матриц смежности экземпляров медицинских технологических процессов в задаче автоматического синтеза обобщенной схемы технологических процессов;
- разработан новый алгоритм построения обобщенной схемы технологического процесса на основе прецедентной информации;
- предложен метод выявления точек ветвления условного маршрута на основе классификации экземпляров технологических процессов;
- разработаны архитектуры и реализованы программные средства автоматического синтеза обобщенных технологических процессов и методы их персонализации.

Практическая значимость. Предложенные методы и подходы к разработке систем поддержки технологических процессов могут быть использованы в крупных медицинских учреждениях для построения систем управления медицинскими технологическими процессами, в университетах для обучения студентов медиков и психологов, в крупных частных компаниях для подбора персонала и оценки психологического состояния сотрудников, в частных компаниях, внедривших системы управления бизнес-процессами, и др.

Апробация результатов исследования. Основные положения работы докладывались и обсуждались на международных научных и научно-практических конференциях: «Интеллектуальный анализ и информационные технологии», «Системный анализ и информационные технологии», EANS congress, «Теория и практика системного анализа», «Информационные технологии в медицине», «Нечеткие системы и мягкие вычисления», «Информационные и телекоммуникационные системы: информационные технологии в научных и образовательных процессах», международная конференция по искусственному интеллекту (КИИ), всероссийская

конференция по проблемам математики, информатики, физики и химии, на семинарах ИСА РАН и Медицинского центра Центрального банка Российской Федерации.

Публикации. Основные результаты, полученные по теме диссертационной работы, опубликованы в 12 печатных работах (в том числе 5 публикаций в ведущих рецензируемых научных изданиях, рекомендованных Высшей аттестационной комиссией, 7 публикаций в трудах научных конференций).

Структура и объем работы. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы и двух приложений. Диссертация содержит 144 страницы, 28 рисунков, 12 таблиц, 146 наименований в списке используемой литературы.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, сформулированы цель и задачи исследования, научная новизна, практическая значимость полученных результатов, а также приведены данные о структуре и объеме диссертации.

В первой главе приведено описание технологического процесса (параграф 1.1). В параграфе 1.2 указаны особенности технологических процессов в медицине и психологии. Параграф 1.3 посвящен обзору методов моделирования и поддержки технологических процессов. Обзор включает в себя компьютерные технологии, применяемые для поддержки медицинских технологических процессов, методы и алгоритмы автоматического синтеза технологических процессов, информационные технологии, применяемые в психологии.

Вторая глава посвящена описанию модели технологического процесса на основе операторной теории.

Моделью технологического процесса является следующая семерка $M = \langle I, CON, IMP, O, CON^I, IMP^I, O^I \rangle^1$, где

- I – множество с заданным на нём отношением линейного порядка (например, множество натуральных чисел) \leq ;
- CON – множество ограничений (условий выполнения операторов);
- IMP – множество императивов;
- O – множество операторов;
- CON^I – отображение, которое каждому элементу из множества I ставит в соответствии элемент из множества CON , т.е. $CON^I : I \rightarrow CON$;
- IMP^I – отображение, которое каждому элементу из множества I ставит в соответствии элемент из множества IMP , т.е. $IMP^I : I \rightarrow IMP$;
- O^I – отображение, которое декартову произведению $I \times I$ ставит в соответствии элемент из множества O , т.е. $O^I : Int \rightarrow O$, где $Int \subseteq I \times I$ такие что,

$\forall t \in I \quad CON^I(t) = Con(t)$, где, а $Con(t)$ – элемент CON ,

$\forall t \in I \quad IMP^I(t) = Imp(t)$, где $Imp(t)$ – элемент IMP ,

$\forall (t_1, t_2) \in Int \quad O^I(t_1, t_2) = o(t_1, t_2, T_t, V_t, Loc_t)$, где $o(t_1, t_2, T_t, V_t, Loc_t)$ – элемент O .

В медицинских технологических процессах в качестве ограничений выступают требования на продолжительность этапов МТП, наиболее важные показатели, требующие объективного исследования, критериальное подтверждение диагноза (если таковое существует), критерии перевода в другое отделение, критерии смены терапии, критерии выписки из стационара,

¹ Назаренко Г.И., Осипов Г.С. Основы теории медицинских технологических процессов. Ч. 1. – М.: ФИЗМАТЛИТ, 2005. – 144 с.

условия применения лечебного мероприятия и др. К императивам относятся указания по сбору данных из всех допустимых источников, о проведении исследований и о выборе специалистов.

Каждый оператор $o \in O$ имеет следующий вид:

$$o = o(t_1, t_2, T, V, Ex),$$

где t_1, t_2 – время начала и окончания выполнения оператора, T – периодичность, Ex – отметка о том, что оператор обязательно должен быть выполнен до конца, или его выполнение можно прервать, V – объем мероприятий.

Если ограничения и императивы описывают состояния, то операторы изменяют состояния, воздействуя на них.

Изменения состояний происходят путем изменения значений признаков с помощью вычисления некоторых функций, если значения признака носят количественный характер; удаления старого признака и добавлении нового, если значения признаков носят неколичественный характер; добавления и удаления концептуальных элементов. В медицинских технологических процессах в качестве концептуальных элементов выступают патологические процессы.

Таким образом, каждый оператор $o(t_1, t_2, T, V, Ex)$ имеет вид:

$$\left\langle \begin{array}{l} \Phi \\ P_1 \\ P_2 \\ Z_1 \\ Z_2 \end{array} \right\rangle,$$

где Φ - семейство функций времени из множества $F = \{f_1, f_2, \dots, f_n\}$, действующих из некоторых декартовых произведений областей значений признаков $X_{i1}^T, X_{i2}^T, \dots, X_{ik}^T$ в области X_j^T – множества значений признаков; P_1 – множество новых появляющихся признаков, задаваемых экстенсиональным образом (множествами значений); P_2 – множество исчезающих признаков, задаваемых экстенсиональным образом; Z_1 – множество новых патологических процессов; Z_2 – множество исчезающих патологических процессов.

Определение 2.1. (Г.С. Осипов, Г.И. Назаренко). Оператор o будем называть применимым к состоянию $State(t)$, если его условие применимости C таково, что $C \subseteq State(t)$.

Определение 2.2. (Г.С. Осипов, Г.И. Назаренко). Последовательным применением операторов o_1 и o_2 называется оператор

$$o_3 = o_1 o_2 = \left\langle \begin{array}{l} \Phi^3 \\ P_1^3 \\ P_2^3 \\ Z_1^3 \\ Z_2^3 \end{array} \right\rangle,$$

такой, что $\Phi^3 = 1 \cup \Phi^1 \cup \Phi^2 \cup \Phi^2 \bullet \Phi^1$, где 1 – единичный оператор, Φ^1 и Φ^2 – компоненты операторов o_1 и o_2 соответственно, $\Phi^2 \bullet \Phi^1$ – композиция этих компонент, $P_1^3 = P_1^1 \cup P_1^2$, $P_2^3 = P_2^1 \cup P_2^2$, $Z_1^3 = Z_1^1 \cup Z_1^2$, $Z_2^3 = Z_2^1 \cup Z_2^2$.

Определение 2.3. (А.И. Молодченков). Последовательным применением n операторов $o_1 \dots o_n$, называется оператор

$$\tilde{o} = o_1 o_2 \dots o_n = \begin{Bmatrix} \Phi^{\sim} \\ P_1^{\sim} \\ P_2^{\sim} \\ Z_1^{\sim} \\ Z_2^{\sim} \end{Bmatrix}$$

такой, что $\Phi^{\sim} = 1 \cup \Phi^1 \cup \Phi^2 \cup \dots \Phi^n \cup \bigcup_{k=2}^n \text{Com}(k)$, где 1 – единичный оператор, Φ^1 , $\Phi^2 \dots \Phi^n$ – функциональные компоненты операторов o_1 , $o_2 \dots o_n$ соответственно, $\text{Com}(k) = \bigcup_{i=1}^{C_n^k} \bullet \text{com}_i$, где $\bullet \text{com}_i$ – композиция элементов com_i , com_i – i -ое k -элементное сочетание множества $\{\Phi_n, \Phi_{n-1}, \dots, \Phi_1\}$ такое, что все элементы $\Phi^{i_1}, \Phi^{i_2}, \dots, \Phi^{i_k} \in \text{com}_i$ ($i_1, i_2, \dots, i_k \in \mathbb{N}$) упорядочены по убыванию значений i_1, i_2, \dots, i_k , $P_1^{\sim} = P_1^1 \cup P_1^2 \cup \dots \cup P_1^n$, $P_2^{\sim} = P_2^1 \cup P_2^2 \cup \dots \cup P_2^n$, $Z_1^{\sim} = Z_1^1 \cup Z_1^2 \cup \dots \cup Z_1^n$, $Z_2^{\sim} = Z_2^1 \cup Z_2^2 \cup \dots \cup Z_2^n$.

В конце первого параграфа автором данной работы были доказаны следующие леммы:

Лемма 2.1 (Г.С. Осипов, Г.И. Назаренко). Последовательное применение операторов ассоциативно, т.е. $(o_1 o_2) o_3 = o_1 (o_2 o_3)$.

Лемма 2.2 (Г.С. Осипов, Г.И. Назаренко). Последовательное применение операторов не коммутативно, т.е. $o_1 o_2 \neq o_2 o_1$.

В параграфе 2.2 описаны элементы технологического процесса и маршруты в нем. Обозначим множество элементов через $E = \{e_1, e_2, \dots, e_n\}$, где каждый $e_i = \langle \text{Con}(t_i), \text{Imp}(t_i), O(t_i, t_{i+1}, T_i, V_i, \text{Loc}_i, \text{Ex}_i) \rangle$. Каждому этапу технологического процесса соответствует несколько элементов, т.е. k -й этап представляет собой подмножество E' декартова произведения множеств $\text{CON}(t_k)$, $\text{IMP}(t_k)$ и $O(t_k, t_{k+1}, T_k, V_k, \text{Loc}_k, \text{Ex}_k)$:

$$E' \subseteq \text{Con}(t_k) \times \text{Imp}(t_k) \times O(t_k, t_{k+1}, T_k, V_k, \text{Loc}_k, \text{Ex}_k).$$

Однако при такой записи не ясно, в какой последовательности выполняются операторы из множества O . Стоит отметить, что с операторами из множества O могут быть связаны условия их выполнения, что приводит к появлению точек ветвления технологического процесса. Элементы из множества E образуют маршруты технологического процесса, а процесс построения этих маршрутов называется маршрутизацией. Для снятия неопределенности выделяют четыре вида маршрутизации: последовательная,

параллельная, условная и итеративная. Последовательная маршрутизация соответствует временному отношению строгого линейного порядка на множестве элементов. Это означает, что мероприятие $e(t_2) \in E$ может быть выполнено *только* после того, как будет выполнено $e(t_1) \in E$. При параллельной маршрутизации порядок выполнения мероприятий не имеет значения. При условной маршрутизации производится выбор между двумя или более маршрутами. Итеративная маршрутизация соответствует повторению некоторого маршрута, пока не будут выполнены критерии выхода из итерации.

В параграфе 2.3 представлено графическое представление технологического процесса. Последовательное расположение стрелок указывает на строго последовательные лечебные мероприятия (выполняются в определенной последовательности), две стрелки, исходящие из одного прямоугольника или эллипса свидетельствуют о параллельной маршрутизации (не важна последовательность выполнения лечебных мероприятий), две стрелки, исходящие из ромба означают условную маршрутизацию (производится выбор между двумя или более маршрутами), а стрелка, исходящая из ромба и входящая в некоторую вершину, предшествующую исходящей, обозначает итеративную маршрутизацию (повторение некоторого маршрута пока не будут выполнены критерии выхода из итерации). На рисунке 2.1 представлено графическое представление различных типов маршрутов в потоках работ.

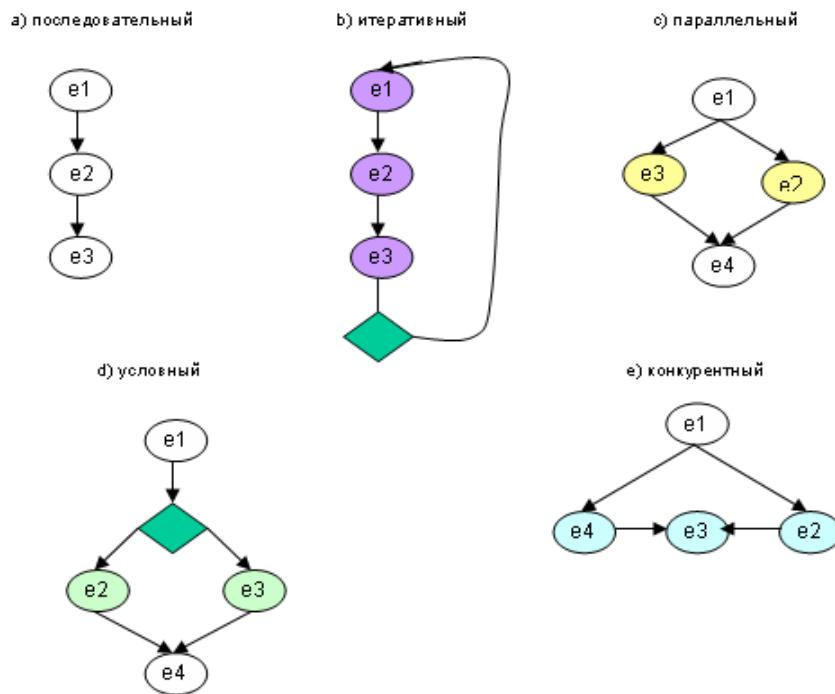


Рисунок 2.1. Типы маршрутов в потоках работ

Медицинский технологический процесс должен соответствовать стандарту лечения нозологии или протоколу ведения больных (Российский аналог клинических руководств). Для выполнения автоматической проверки этого соответствия стандарт или протокол ведения больных должны быть представлены в формальном виде. В параграфе 2.4 описан онтологический подход для описания стандартов лечения и протоколов ведения больных.

Онтология – это формальное описание медицинских знаний, касающихся медицинских технологических процессов. В качестве способа представления знаний используется семантическая сеть. Узлы сети соответствуют классам, экземплярам классов и другим элементам медицинского технологического процесса. Дуги семантической сети соответствуют отношениям между ними.

Для описания онтологии был использован язык представления онтологий OWL (Ontology Web Language) в диалекте DL (Description Logic) и редактор онтологий Protégé (<http://protege.stanford.edu/>). Основными элементами онтологии являются классы, их экземпляры, слоты (представляющие свойства классов и экземпляров) и фасеты, задающие дополнительную информацию о слотах. Классы обладают двумя типами свойств: свойства-данные и свойства-отношения. Свойства-данные соответствуют характеристикам, которым можно присвоить некоторое значение (числовое, строковое и т.п.). Например, имя, идентификатор, возраст. Свойства-отношения представляют собой связи с классами, экземплярами классов и другими элементами онтологии.

Класс «Медицинский_технологический_процесс» содержит множество лечебно-диагностических мероприятий (действий), применимых к пациенту, и множество этапов, из которых состоит МТП. Обязательно присутствует начальное действие (start_action), которое является точкой входа МТП. Класс «Этап_ТП» является подклассом класса «Медицинский_технологический_процесс» ввиду того, что каждый этап медицинского технологического процесса – это его подпроцесс. Этап МТП может быть разбит на множество более мелких этапов и т.д.

Действия в МТП носят лечебно-диагностический (например, «сбор анамнеза и жалоб при заболеваниях легких и бронхов») или административный (например, «перевод в клиническое отделение») характер, связаны с временными характеристиками, местом проведения, исполнителем и т.д. Действия могут быть простыми и составными. Примером простого действия является «Пульсоксиметрия». Составное действие состоит из набора простых действий и может включать в себя составные мероприятия. Например, «Диагностические исследования в приемном отделении» при обострении бронхиальной астмы включает в себя следующие действия: «Биохимический анализ крови», «Общий анализ крови», «Общий анализ мочи», «Пульсоксиметрия» и др.

Класс времени делится на три группы. Первая группа касается временных периодов, таких как «Время пребывания в стационаре», «Первый день после проведения хирургического лечения» и т.д. Вторая группа относится к временным моментам и связана с моментами начала и завершения лечебного мероприятия. Третья группа касается временных отношений между действиями. Эта группа может быть разделена на две категории. Первая категория называется относительным временным отношением, которое описывает положение выполнения двух вмешательств относительно друг друга. Например, действие b выполняется через 4 часа после выполнения действия a. Вторая категория называется отношением временного интервала, которое сопоставляет некоторый известный период времени с интервалом между

началом и окончанием выполнения двух действий. Эта категория служит для выявления отклонений в процессе исполнения МТП. В качестве известного интервала выступает временной интервал первой категории между двумя действиями, указанный в онтологии. Далее происходит вычисление времени между началом и окончанием выполнения этих действий и сопоставление. В связи с этим для действий допускается указывать по два момента времени допустимого начала и окончания выполнения.

С каждым действием, кроме времени, связаны «Роль», «Место_проведения» и «Ресурсы». Эти классы соответствуют различным механизмам поддержки выполнения лечебно-диагностических мероприятий. Класс «Роль» имеет два подкласса: «Пациенты» и «Персонал». «Персонал» делится на следующие подклассы: «Врачи», «Медсестры», «Обслуживающий_персонал» и «Руководство». Классу «Место_проведения» соответствуют подразделения медицинского учреждения, в которых выполняются соответствующие действия. В качестве ресурсов выступают оборудование, медикаменты и т.п.

Каждому из типов маршрутов в онтологии соответствуют свои классы и отношения. Для обозначения того, что два действия e_1 и e_2 выполняются последовательно, было введено отношение $next(e_1, e_2)$, которое означает, что действие e_2 выполняется только после выполнения e_1 . Для остальных типов маршрутов были введены одноименные классы. Для тех случаев, когда время начала выполнения $e_2 \leq$ времени окончания выполнения e_1 , было введено понятие конкурентного маршрута. Класс «Параллельный_маршрут» является подклассом класса «Маршрутизация» и обладает свойством-отношением «action», которое связывает экземпляр класса «Параллельный_маршрут» с экземплярами класса «Элемент_техпроцесса». Это отношения типа «один ко многим», т.к. в одном параллельном маршруте может быть исполнено несколько элементов технологического процесса. Класс «Условный_маршрут» также является подклассом класса «Маршрутизация» и обладает свойством-данным условие и двумя свойствами-отношениями «true_action» и «false_action». Свойство-данное принимает строковое значение, в котором в определенном формате описывается условие дальнейшего движения по технологическому процессу. На основании описанных условий формируется система правил, позволяющая управлять МТП в процессе исполнения. Свойства «true_action» и «false_action» аналогичны свойству «action» класса «Параллельный_маршрут» и соответствуют дальнейшему движению по технологическому процессу, если условие истинно или ложно. Класс «Итеративный_маршрут» является подклассом «Маршрутизация» и обладает свойством-данным условие и свойством-отношением «action». В свойстве условие прописывается условие выхода из итерации. Свойство «action» аналогично соответствующему свойству в классе «Параллельный_маршрут». Кроме классов, соответствующих различным типам маршрутов, к элементам технологического процесса относятся лечебно-диагностические и организационные мероприятия. Соответствующие классы введены как

подклассы класса «Элемент_техпроцесса». Далее лечебно-диагностические и организационные мероприятия делятся на более мелкие классы и т.д.

В конце параграфа приведен пример онтологии процесса лечения бронхиальной астмы.

Третья глава посвящена описанию алгоритма автоматического построения общего описания технологического процесса на основе прецедентной информации.

В параграфе 3.1 описан алгоритм автоматического синтеза технологического процесса на основе прецедентной информации. Через $\Omega(O)$ обозначим семейство последовательностей вида $w = \langle o_i, o_j, \dots, o_k \rangle$, где w – упорядоченное множество операторов, $i < j < \dots < k$ – элементы множества натуральных чисел N . На множестве $\Omega(O)$ задано некоторое отношение эквивалентности ρ , порождающее фактор-множество Ω_ρ и разбивающее множество всех примеров на классы эквивалентности (например, по нозологическим формам). Ближайшей целью является построение полного описания всех классов эквивалентности. Другими словами, задача состоит в том, чтобы опираясь на множество примеров, входящих в классы эквивалентности $\{w\}$, построить такие описания $G(\{w\})$, что каждый пример, входящий в $\{w\}$, удовлетворял бы его описанию $G(\{w\})$.

Описание классов эквивалентности будем строить в виде графа. То обстоятельство, что все элементы в $\{w\}$ удовлетворяют описанию, означает, что для каждого примера w из $\{w\}$ в графе $G(\{w\})$ должен найтись эквивалентный ему подграф. Все графы $G(\{w\})$ будем представлять посредством их матриц смежности. Введена операция покомпонентного сложения матриц смежности: если $A = [a_{ij}]$ и $B = [b_{ij}]$ – матрицы смежности, то $A + B = [a_{ij}] + [b_{ij}] = \max([a_{ij}], [b_{ij}])$.

Алгоритм 3.1 (общий алгоритм автоматического построения ТП)

- 1) построение матриц смежности, описывающих частные технологические процессы (экземпляры);
- 2) разбиение частных технологических процессов на классы эквивалентности и их кластеризация;
- 3) выбор опорного примера в каждом классе эквивалентности;
- 4) синтез обобщенной схемы технологического процесса для каждого класса эквивалентности путем объединения его экземпляров.

Алгоритм 3.2 (алгоритм построения матриц смежности, описывающих частные технологические процессы):

- 1) извлечение действий (операторов);
- 2) сортировка списков действия по времени их выполнения (построение $\Omega(O)$);
- 3) построение матриц смежности для каждого примера, входящего в $\{w\}$.

Все действия, которые были выполнены в частном ТП, извлекаются из лог файла, базы данных и других источников, и записываются в массив w_i , где i – индекс примера технологического процесса. После извлечения всех действий,

они сортируются по возрастанию относительно времени их выполнения. В результате получится множество экземпляров $\Omega(O)$.

Следующим шагом является построение описаний в виде матриц смежности каждого экземпляра в $\Omega(O)$. С номером каждой строки и столбца матрицы смежности связан некоторый оператор (лечебно-диагностическое мероприятие). С одним и тем же оператором может быть связано несколько строк, т.к. при последовательной маршрутизации одно и то же лечебное мероприятие могут делать несколько раз и отсутствует итеративный маршрут.

Пусть $M^k = [m_{ij}^k]$ – матрица смежности для k -го примера технологического процесса, h -размерность M^k , (перед началом работы алгоритма построения матрицы смежности для частного технологического процесса $h = 0$).

Алгоритм 3.3 (алгоритм построения матрицы смежности для экземпляра технологического процесса):

- 1) считывается оператор o_i из экземпляра $x \in \Omega(O)$, где $i = 0, \dots, n$ – индекс оператора в экземпляре технологического процесса;
- 2) если этот оператор ранее не выполнялся, то он добавляется в матрицу M^k ; $M^k[h, h+1] = 1$, во всех остальных пустых ячейках устанавливаются 0, h и i увеличиваются на 1, переход к шагу 1;
- 3) если оператор ранее выполнялся и находится в j -ом столбце ($j < i$), то запускается процедура поиска итерации;
- 4) если удалось обнаружить итерацию, то $M^k[h, j] = 1$, запуск процедуры объединения операторов, расположенных между строками j и h матрицы M^k и после i -го оператора в x , i увеличивается на количество элементов в итерации, переход к шагу 1;
- 5) если итерации нет, то оператор добавляется в матрицу M^k , $M^k[h, h+1] = 1$, во всех остальных пустых ячейках устанавливаются 0, h и i увеличиваются на 1, переход к шагу 1.

Итеративные маршруты возникают в том случае, если при анализе примера технологического процесса обнаружен оператор, выполненный ранее. В этом случае проверяется, совпадают ли множества операторов, выполненных вслед за предыдущим и текущим выполнением обнаруженного оператора. Итеративный маршрут порождается, если совпадение обнаружено. Для определения совпадения используется следующая функция релевантности:

$$rel(\pi_1, \pi_2) = \varphi(\rho_1(e_1, \pi_2), \dots, \rho_n(e_n, \pi_2)),$$

где π_1, π_2 – множества операторов между предыдущим и текущим положением оператора и следующим за текущим, e_1, \dots, e_n – элементы π_1 . Положим

$$\rho(e_i, \pi_2) = \begin{cases} 1, & \text{если } e_i \in \pi_2 \\ 0, & \text{если } e_i \notin \pi_2 \end{cases},$$

а φ – линейная комбинация ρ_i , где $i \in \{1, \dots, n\}$, нормированная на 1, т.е.

$$rel(\pi_1, \pi_2) = \frac{1}{n} \sum_{i=1}^n \rho_i.$$

Если значение релевантности больше некоторого коэффициента k , то будем считать, что итерация обнаружена, иначе нет. Значение k определяется в зависимости от решаемой задачи.

Для разбиения экземпляров технологических процессов на классы эквивалентности на множестве $\Omega(O)$ вводится отношение эквивалентности ρ . Для медицинских технологических процессов таким отношением является выбор примеров по общей нозологии. В один класс эквивалентности записываются подпроцессы, направленные на лечении одного заболевания.

Для построения классов эквивалентности введем следующее отношение:

$$f : \Omega(O) \rightarrow U,$$

где U – множество нозологических форм.

Два экземпляра $x \in \Omega(O)$ и $y \in \Omega(O)$ принадлежат одному классу эквивалентности тогда и только тогда, когда $f(x) = f(y)$. Алгоритм разбиения множества $\Omega(O)$ на классы эквивалентности проиллюстрирован на рисунке 3.3.

Для каждого экземпляра $y \in \Omega(O)$ производится проверка на равенство $f(x) = f(y)$ со всеми x – примерами, входящими в классы эквивалентности $\{w\}$. Если удалось найти такой пример x , такой что $f(x) = f(y)$, то записываем y в класс эквивалентности, которому принадлежит x . Если не удалось найти ни одного примера x , что $f(x) = f(y)$, то добавляем во множество $\{w\}$ новый класс эквивалентности и записываем в него y .

После разбиения частных технологических процессов на классы эквивалентности производится синтез обобщенной схемы ТП. Пусть $\{M^k\}$ – множество матриц смежности, описывающих частные технологические процессы для k -го класса эквивалентности.

Алгоритм 3.4 (алгоритм автоматического синтеза обобщенной схемы технологических процессов):

- 1) приведение всех матриц смежности из $\{M^k\}$ к одной размерности таким образом, чтобы их строки и столбцы указывали на одни и те же операторы;
- 2) построить матрицу M – являющуюся объединением всех матриц из $\{M^k\}$;
- 3) убрать шумы в получившемся описании обобщенной схемы технологического процесса;
- 4) выявить все маршруты в описании обобщенной схемы технологического процесса;
- 5) построить условия в точках ветвления технологического процесса.

На первом шаге в матрицах из $\{M^k\}$ несколько раз производятся операции добавления и перестановки строк и столбцов до выполнения описанных выше условий. Матрица M строится путем объединения матриц смежности из $\{M^k\}$.

После построения матрицы M необходимо убрать шумы в описании технологического процесса. Через $a \rightarrow b$ обозначим тот факт, что оператор b выполняется после выполнения оператора a . Для каждой пары операторов (a, b) в матрице M , которые находятся в отношении $a \rightarrow b$, рассчитывается значение следующей метрики ¹:

$$p = \frac{e^{LR}}{1 + e^{LR}},$$

где значение LR , рассчитывается по следующей формуле:

$$LR = -8.280 + 6.376 * LM + 4.324 * GM + 8.654 * CM$$

Метрика LM (local metric) ²:

$$LM = P - 1.96 \sqrt{\frac{P(1-P)}{N+1}},$$

$$\text{где } P = \frac{(a \rightarrow b)}{N+1}, \quad N = (a \rightarrow b) + (b \rightarrow a).$$

Метрика GM (global metric) ³:

$$GM = ((a \rightarrow b) - (b \rightarrow a)) \frac{\#L}{\#A * \#B}$$

Метрика CM (causal metric):

$$CM = \frac{C(a \rightarrow b)}{\#A},$$

где значение $C(a \rightarrow b)$ рассчитывается следующим образом; для каждого прецедента $C(a \rightarrow b)$ увеличивается на величину δ^n . В результате проведенных экспериментов авторами этой метрики было установлено значение δ равным 0.8. n – количество элементов между a и b . В случае если элемент b следует сразу после a (т.е. $n = 0$), то $C(a \rightarrow b)$ увеличивается на 1.

Пороговое значение σ для метрики p было получено экспериментальным путем и равно 0.8.

После работы алгоритма отсечения шумов из матрицы смежности удаляются все элементы, не имеющие входных связей и не являющиеся начальными, и элементы, не имеющие выходных связей и не являющиеся конечными.

¹ T. Weijters, W.M.P. van der Aalst. Process Mining: Discovering Workflow Models from Event-Based Data. In Kröse, B. et. al, (eds.): Proceedings 13th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC'01), 25-26 October 2001, Amsterdam, The Netherlands, pp. 283-290.

² Mărușter, L., Weijters, A.J.M.M., Aalst, W.M.P., and Bosch, A. 2002. Process mining: Discovering direct successors in process logs. In S. Lange, K. Satoh, and C.H. Smith (Eds.), Proceedings of the 5th International Conference on Discovery Science (Discovery Science 2002), Berlin: Springer-Verlag, vol. 2534: pp. 364–373.

³ Mărușter L. et al. A rule-based approach for process discovery: Dealing with noise and imbalance in process logs //Data mining and knowledge discovery. – 2006. – Т. 13. – №. 1. – С. 67-87.

Завершающим этапом построения обобщенной схемы технологических процессов является извлечение различных типов маршрутов (потоков работ) на основе извлеченной прецедентной информации. Извлечение различных типов маршрутов в ТП производится путем применения к матрице M системы правил.

Параграф 3.2 посвящен описанию решения задачи персонализации медицинских технологических процессов. Пусть имеется вновь поступивший в медицинское учреждение пациент Z . Задача персонализации заключается в следующем:

- на основе схем лечения K пациентов строится общая схема лечения конкретной нозологической формы всех больных с близкими к пациенту Z состояниями;
- для вновь поступившего пациента Z , взяв за основу построенную обобщенную схему и используя его индивидуальные особенности, построить схему лечения для пациента Z .

Работа предложенного алгоритма задачи персонализации продемонстрирована на примере диагностики и лечения пациентов с болью в спине. Для выполнения адаптации обобщенной схемы медицинского технологического процесса под индивидуальные особенности пациентов необходимо построить условия выбора маршрута в точках ветвления.

Эту задачу можно сформулировать следующим образом: даны обобщенная схема технологического процесса и множество примеров частных технологических процессов с описаниями состояний пациентов в каждый момент времени исполнения ТП, необходимо в точках ветвления (в условных и итеративных маршрутах) определить условия выбора маршрута в ТП. Выбор маршрута можно записать в виде правил, а решение задачи построения этих правил сводится к решению задачи классификации.

В этом разделе описывается применение алгоритма последовательного порождения правил AQ¹ для решения задачи классификации и проводится оценка результатов его работы. AQ-алгоритм является достаточно простым, не использует статистические функции, не зависит от типов входных данных, показывает хорошие результаты на небольших выборках и дает полное и непротиворечивое описание рассматриваемой ситуации. Недостатком его является сильная зависимость от примеров, с которыми он работает, и, в некоторых случаях, получение довольно громоздких правил.

AQ-обучение основывается на идее постепенного покрытия обучающих данных с помощью последовательно порождаемых правил и используется целым семейством алгоритмов. Суть метода заключается в поиске множества правил (конъюнкций пар атрибут-значение или, в общем случае, произвольных предикатов), которое покрывает все положительные (+) примеры и ни одного отрицательного (–) примера. Вместо разбиения множества примеров AQ-алгоритм шаг за шагом обобщает описания выбранных положительных примеров, называемых опорными примерами.

¹ Молодченков А.И. Применение AQ-алгоритма для персонификации лечебно-диагностических процессов //Теория и практика системного анализа: Труды I Всероссийской научной конференции молодых ученых. – Т. I. – Рыбинск: РГАТА им.П.А. Соловьева, 2010, с.79-84.

В обучающиеся системы, построенные на основе алгоритма AQ, можно интегрировать любые базовые знания, т.к. они часто могут быть представлены в виде системы правил. Более того, AQ не накладывает ограничения на типы знаний. Они могут быть числовыми, интервальными, символьными и др. При построении правил выбора маршрута в точках ветвления AQ-алгоритм использует дополнительные знания из стандартов лечения и протоколов ведения больных, описанные в онтологической форме. Работа AQ-алгоритма представлена на примере определения степени обострения бронхиальной астмы.

Приведен алгоритм решения задачи определения условий выбора направления движения в условных и итеративных маршрутах. Этот алгоритм состоит из следующих шагов:

- 1) для каждой точки ветвления $p \in P$ в массив paths записываются все пути, ведущие из p до точки их слияния, states – все состояния пациентов, в момент нахождения процесса лечения пациента в точке, $rule^p = 0$;
- 2) если в paths нет необработанных маршрутов, то в Rules добавляется $rule^p$, переход к шагу 1;
- 3) иначе path= необработанный маршрут из paths, PPaths – все частные ТП из $\{M^k\}$, включающие в себя path, в PosStates записываются все состояния из states, связанные с частными ТП из PPaths, в NegStates записываются остальные состояния из states;
- 4) в $rule^p$ добавляется результат работы алгоритма AQ, для которого PosStates – положительные примеры, NegStates – отрицательные, переход к шагу 2;

Алгоритм завершит свою работу, когда все точки ветвления $p \in P$ будут обработаны. В результате работы алгоритма во всех точках ветвления в условных и итеративных маршрутах будут определены правила, определяющие условия выбора пути движения по маршруту технологического процесса.

Спецификой медицинских задач является то, что медицинские данные не только разнородны, но и, как правило, не имеют представительной выборки. При решении задачи автоматического построения условий выбора направления движения в точках ветвления маршрутов, основным требованием является то, что построенные правила должны быть очень хорошего качества даже в случаях, когда выборки очень малы. В конце параграфа проведена оценка результатов работы AQ-алгоритма в сравнении с алгоритмами классификации BayesNet (Байесовский классификатор), J48graft (алгоритм C4.5, использующий деревья решений) и JRip (JRipper).

Оценка результатов проводилась путем расчета следующих метрик: чувствительность (Sensitivity), специфичность (Specificity), F-мера, точность (Precision), достоверность (Accuracy), - на трех наборах данных. Два набора были взяты из базы данных UCI¹, предназначеннной для оценки работы методов машинного обучения. Третий набор данных был предоставлен Медицинским

¹ UCI Machine Learning Repository. <http://archive.ics.uci.edu/ml/>

центром Центрального банка Российской Федерации. Для проведения экспериментов была взята реализация AQ-алгоритма AQ21¹. Для алгоритмов BayesNet, J48graft и JRip была взята их реализация в библиотеке WEKA².

Эксперименты показали, что правила, построенные AQ-алгоритмом, по качеству сопоставимы или превосходят результаты указанных выше алгоритмов.

В параграфе 3.3 описаны результаты экспериментальных исследований предложенного в работе алгоритма автоматического синтеза медицинских технологических процессов на основе прецедентной информации.

В отличие от анализа результатов алгоритмов классификации метод k-fold cross validation в чистом виде не подходит для анализа результатов алгоритма автоматического синтеза технологических процессов. Это связано с тем, что прецеденты не содержат положительных и отрицательных примеров. Каждый прецедент описывает поведение технологического процесса в конкретном случае и всегда будет положительным. Некоторые исследователи предлагают искусственно вводить «неправильные» примеры. Но этот способ не подходит при оценке реальных процессов.

Для оценки результатов используется следующая метрика³:

$$f = \frac{1}{2} \left(1 - \frac{m}{c} \right) + \frac{1}{2} \left(1 - \frac{r}{p} \right),$$

где m – количество пропущенных в ходе проверки примера позиций, c – количество позиций, в которые переходила метка, r – количество позиций в которых метка осталась после полной проверки примера, p – количество позиций, в которые метка должна перейти. При $m=r=0$ фитнес функция равна 1.

Физический смысл этой фитнес функции заключается в том, что она оценивает наличие маршрута, в примере из тестовой выборки, в построенной модели технологического процесса. Если ее значение равно 1, это означает, что искомый маршрут полностью имеется в модели технологического процесса. Если значение фитнес функции меньше 1, то не все части искомого маршрута имеются в построенной модели. Если значение ее значение равно 0, то искомый маршрут отсутствует в построенной модели.

Пусть N – число примеров в тестовой выборке, σ_i – i -й пример из тестовой выборки. Тогда фитнес функция для оценки результатов алгоритма автоматического синтеза обобщенной схемы технологического процесса рассчитывается по следующей формуле:

$$F = \frac{1}{2} \left(1 - \frac{\sum_{\sigma_i} m_{\sigma_i}}{\sum_{\sigma_i} c_{\sigma_i}} \right) + \frac{1}{2} \left(1 - \frac{\sum_{\sigma_i} r_{\sigma_i}}{\sum_{\sigma_i} p_{\sigma_i}} \right)$$

¹ JanuszWojtusiak, Ryszard S. Michalski, Kenneth A. Kaufman, JaroslawPietrzykowski The AQ21 Natural Induction Program for Pattern Discovery: Initial Version and its Novel Features / International Conference on Tools with Artificial Intelligence - ICTAI, pp. 523-526, 2006.

² Ian H. Witten, Eibe Frank, Len Trigg, Mark Hall, Geoffrey Holmes, and Sally Jo Cunningham (1999). Weka: Practical machine learning tools and techniques with Java implementations.

³ Wil M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer Publishing Company, Incorporated, 2011, 352 p.

При проведении экспериментов результаты алгоритма, предложенного в этой работе и реализованного на языке C#, сравнивались с результатами α -алгоритма, реализованного в пакете ProM версии 6.3. Эксперименты проводились на трех выборках логов событий. Все эти выборки были разделены на десять подмножеств примеров. Девять подмножество использовались для построения обобщенной модели технологических процессов; десятое подмножество использовалось для оценки результатов работы алгоритмов.

Первая выборка входит в пакет ProM6.3 и применяется для проверки алгоритмов автоматического синтеза технологических процессов на основе прецедентной информации. Выборка в этом примере очень хорошего качества, создана искусственно и без шумов. Выборка содержит 1104 частных процессов, содержащих более 7000 событий.

Вторая выборка была собрана сотрудниками Технологического университета Эйндховена (Нидерланды) в рамках выполнения проекта CoSeLoG¹. Выборка содержит реальные процессы по подаче заявки на строительство в муниципалитет. В файле содержится 1434 частных процессов, содержащих 8577 событий. Выборка содержит шумы.

Третья выборка была собрана сотрудниками Технологического университета Эйндховена². В этой выборке содержится 274 процесса лечения злокачественных новообразований в области шейки матки, содержащих, в общей сложности, более 1000 событий. Выборка была предоставлена Голландским академическим госпиталем и содержит реальные лечебно-диагностические процессы. Стоит отметить, что эта выборка содержит много шума.

Как показали результаты экспериментальных исследований, алгоритм предложенный в данной работе работает не хуже α -алгоритма на хорошей выборке и превосходит по качеству работы на зашумленных выборках. При проведении экспериментальных исследований необходим было показать, что предложенный в работе алгоритм на зашумленных данных не только работает лучше α -алгоритма, но значение фитнес функции, при оценке результатов, превышает 0.6. В области анализа процессов считается, что алгоритм автоматического построения технологического процесса на основе прецедентной информации работает очень хорошо, если оценка результатов его работы превышает пороговое значение, равное 0.6. Эксперименты показали, что результаты работы предложенного в данной работе алгоритма не только превосходят результаты работы α -алгоритма, но и значение фитнес функции предложенного алгоритма значительно превышает 0.6.

Четвертая глава посвящена описанию архитектур программных систем поддержки технологических процессов в области медицины и психологии.

¹ Buijs, J.C.A.M. (2014) Receipt phase of an environmental permit application process ('WABO'), CoSeLoG project. Eindhoven University of Technology. Dataset. <http://dx.doi.org/10.4121/uuid:a07386a5-7be3-4367-9535-70bc9e77dbe6>

² van Dongen, B.F. (2011) Real-life event logs - Hospital log. Eindhoven University of Technology. Dataset. <http://dx.doi.org/10.4121/uuid:d9769f3d-0ab0-4fb8-803b-0d1120ffcf54>

В параграфе 4.1 описана архитектура системы поддержки медицинских технологических процессов. Система управления МТП состоит из трех основных модулей: модуля создания и тестирования технологических процессов, модуля синхронизации с внешними системами и модуля управления МТП. Модуль создания и тестирования МТП предоставляет экспертам полный функционал для создания медицинских технологических процессов. После того, как технологический процесс лечения заболевания был построен, проходит тестирование с помощью подсистемы тестирования технологических процессов.

Если технологический процесс лечения заболевания удачно прошёл тестирование, то он с помощью подсистемы синхронизации и редактирования добавляется в базу процессов лечения заболеваний. Иначе, этот процесс отправляется на доработку.

Основной задачей модуля синхронизации является реализация программных коннекторов к другим системам, используемым в медицинских технологических процессах. После того, как коннектор реализован и проверен, он внедряется в подсистему управления МТП через подсистему конфигурирования.

Модуль управления медицинскими технологическими процессами является ключевым элементом системы поддержки МТП. Его основными задачами являются: поиск в базе технологических процессов (ТП) походящего процесса лечения заболевания в зависимости от диагноза больного, запуск найденного процесса, контроль его выполнения, анализ результатов его выполнения.

Работа подсистемы управления ТП с текущим пациентом заканчивается при завершении его лечения в стационаре.

Реализация алгоритма управления медицинскими технологическими процессами осуществляется с помощью подсистемы редактирования и проверки процесса управления лечением.

В параграфе 4.2 описана архитектура системы поддержки технологических процессов в области практической психологии. Основным компонентом этой системы является база знаний, которая аккумулирует в себе знания экспертов в себе знания экспертов о предметной области. Эксперты вносят изменения в базу знаний с помощью интерфейса эксперта и системы приобретения знаний. Интерфейс эксперта может быть реализован как Windows или Web приложение. Web-интерфейс эксперта используется в тех случаях, когда эксперты территориально находятся в разных городах или странах. При удаленном построении базы знаний разными экспертами в систему приобретения знаний добавляется возможность объединения баз знаний, построенных разными экспертами. Эксперты могут строить базу знаний как с привлечением инженера по знаниям, так и без него. Это зависит от реализованного алгоритма приобретения знаний от экспертов. Решатель обрабатывает введенную пользователем информацию в соответствии с логикой базы знаний. Интерфейс пользователя служит для ввода данных и отображения результатов работы решателя. Еще с помощью интерфейса эксперта можно

работать с компьютеризированными методиками. Эти методики могут быть реализованы как в самой системе поддержки процессов оказания помощи клиентам, так и просто подключены к ней. Интерфейс пользователя может быть реализован в виде Windows или Web приложения.

Существует несколько способов представления знаний: правила, фреймы, семантические сети, неоднородные семантические сети и др.

В параграфе 4.3 описан способ представления знаний в области практической психологии с помощью неоднородных семантических сетей (НСС). Преимуществом НСС является то, что их можно использовать для формализации знаний в слабоструктурированных областях, к которым относится психология. Отличительной особенностью предлагаемого подхода является использование новых отношений на множестве концептов. Эти отношения используются для построения иерархии на множестве измеряемых объектов диагностики (ОД) и построения дерева ОД; для связи диагностических методик с объектами диагностики, свойствами методик, показателями, измеряемыми с помощью выбранных методик, параметрами клиента и др.

База знаний при использовании НСС представляет собой следующую структуру $KB = \langle S, CL, O, C, CT, P, H \rangle$, где S – множество разделов; CL – множество классов; O – множество объектов; C – множество связей; CT – множество типов связей; P – множество процедур; H – множество состояний базы знаний.

Класс cl – это произвольная пара $\langle N, Ch \rangle$, где N – имя класса, Ch – его характеристика, $Ch = \langle P \rangle$ – множество свойств класса cl , N представляет собой одно или несколько слов в некотором алфавите.

Основными составляющим базы знаний являются объекты, представляющие собой различные понятия предметной области (утверждения, события, факты, результаты наблюдений и т.п.). Объекты являются вершинами семантической сети. Объект представляет собой следующую структуру $o = \langle oN, oCh \rangle$, где oN – имя объекта, $oCh = \{A_1:v_1, \dots, A_n:v_n\}$ – его характеристика. Характеристика объекта задается непустым множеством атрибутов и их значений. Каждому атрибуту A_i (независимо от рассматриваемого объекта) сопоставлен домен (шкала) $D_i = D(A_i)$ – непустая область возможных значений атрибута A_i , значение v_i должно принадлежать D_i . В качестве характеристики объекта могут выступать либо другие объекты, либо атрибуты. Атрибутом называется свойство объекта, имеющее область значений и способное изменяться в зависимости от ситуации в пределах этой области. Большинство объектов характеризуется множеством атрибутов и множествами значений каждого атрибута. Если свойством объекта объявляется другой объект, то к свойствам первого объекта присоединяются все свойства второго. Для каждой конкретной области в базе знаний используется свой набор объектов, их атрибутов, доменов и др.

Объекты в базе знаний сгруппированы в разделы. Объекты могут попадать в разные разделы в зависимости от их функциональной роли. Например, при поддержке процесса психодиагностики объекты делятся на

разделы «Клиент», «Объекты диагностики», «Свойства методик» и др. Имя раздела, как правило, играет большую информативную роль. Имя раздела не должно совпадать с именами объектов. Количество разделов, как и объектов не ограничено.

Для того чтобы наличие одних объектов в предметной области могло повлечь порождение либо исключение других, между объектами в базе знаний устанавливаются связи. Связь можно установить между любыми двумя объектами, хотя наиболее важны связи между признаками и гипотезами.

Исполнение связей происходит путем вызова соответствующих процедур обработки связей.

В параграфе 4.4 описаны основные классы и структуры данных.

Архитектуры систем, использующих результаты работы представлены на рисунках 4.1 и 4.2.

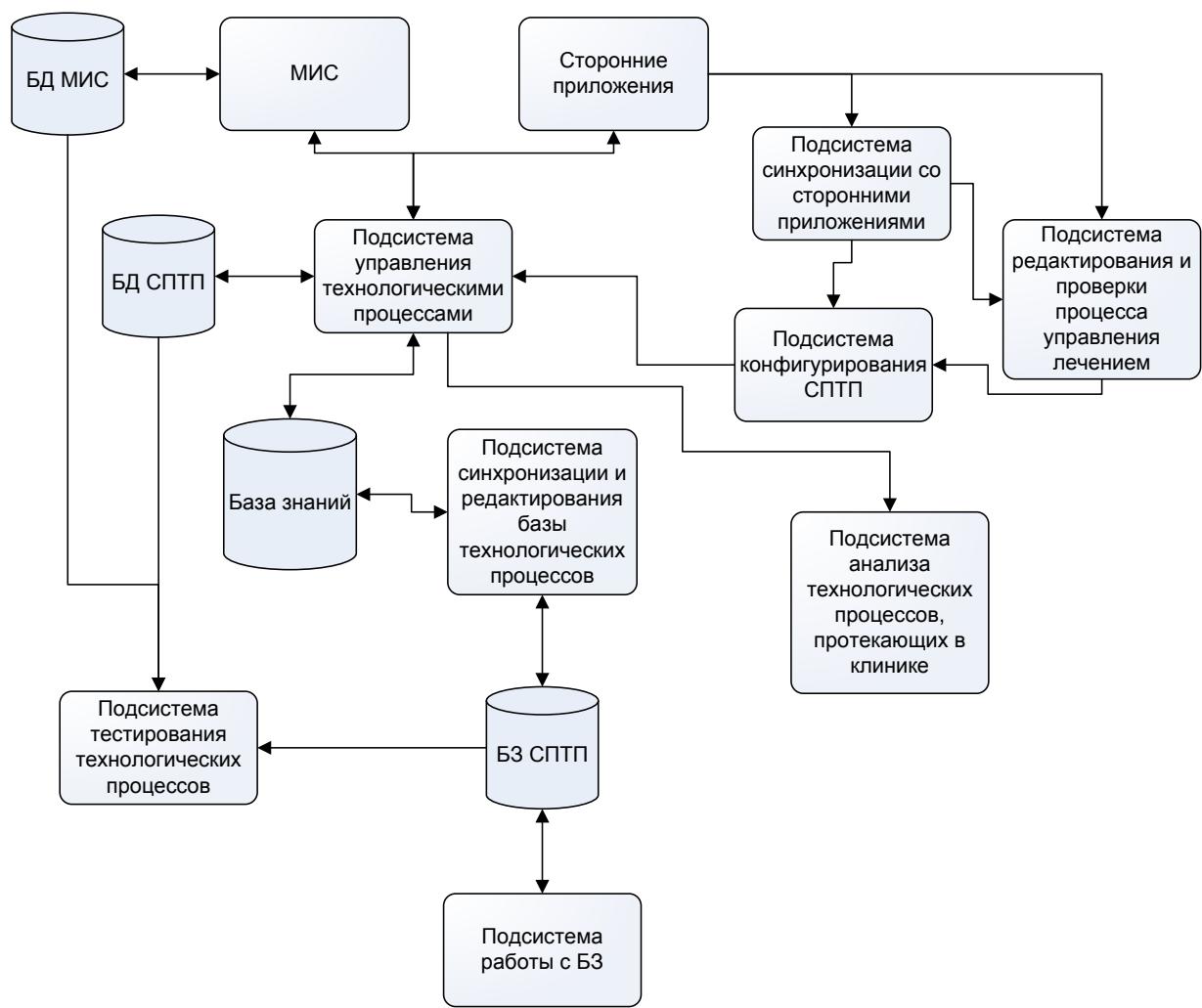


Рисунок 4.1. Архитектура системы поддержки медицинских технологических процессов

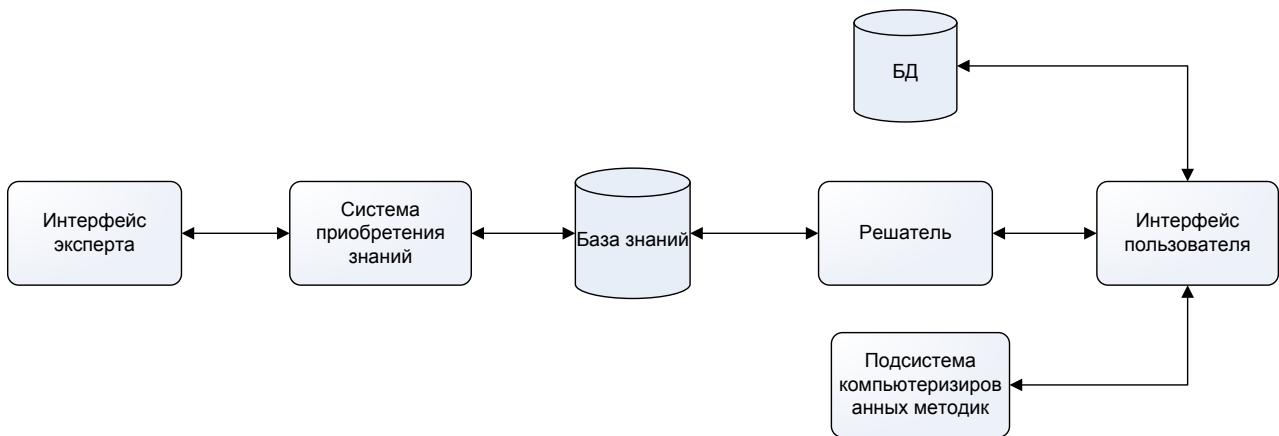


Рисунок 4.2. Архитектура системы поддержки процессов оказания помощи клиентам в области практической психологии

В заключении приведены основные результаты и выводы диссертационной работы.

Результаты, выносимые на защиту:

- впервые предложено матричное представление медицинских технологических процессов;
- доказаны утверждения об ассоциативности и не коммутативности последовательного применения произвольного числа операторов МТП;
- введена операция покомпонентного сложения матриц смежности экземпляров медицинских технологических процессов в задаче автоматического синтеза обобщенной схемы технологических процессов;
- разработан новый алгоритм построения обобщенной схемы технологического процесса на основе прецедентной информации;
- предложен метод выявления точек ветвления условного маршрута на основе классификации экземпляров технологических процессов;
- разработаны архитектуры и реализованы программные средства автоматического синтеза обобщенных технологических процессов и методы их персонализации.

В приложении А описан комплекс экспертных систем «Психология», предназначенный для поддержки процессов психодиагностики и психологического консультирования.

В приложении Б описаны основные составляющие системы поддержки процесса принятия решения о проведении коронарного вмешательства.

В приложении В предложен программный код и описание основных функций программной реализации алгоритмов автоматического синтеза медицинских технологических процессов.

Предложенные в диссертационной работе методы и алгоритмы были применены для разработки комплекса экспертных систем «Психология» и ряда систем поддержки различных этапов лечебно-диагностических процессов в Медицинском центре Банка России.

ПУБЛИКАЦИИ

Основные положения диссертации отражены в следующих публикациях:

Публикации в журналах, входящих в перечень ВАК

1. Молодченков А.И. Автоматический синтез типовых схем лечения на основе медицинских данных. Информационно-измерительные и управляющие системы. - 2009. - Т. 7, №12, с. 90-93.
2. Макаров Д.А., Молодченков А.И. Создание систем приобретения знаний для построения медицинских экспертных систем на основе ядра программных инструментальных средств "MedExp". Информационно-измерительные и управляющие системы. - 2009. - Т. 7, №12, с. 86-89.
3. Г.И. Назаренко, Г.С. Осипов, А.Г. Назаренко, А.И. Молодченков. Интеллектуальные системы в клинической медицине. Синтез плана лечения на основе прецедентов // Информационные технологии и вычислительные системы. № 1, 2010, стр. 24-35.
4. Назаренко Г.И., Клейменова Е.Б., Яшина Л.П., Молодченков А.И., Пающик С.А., Константинова М.А., Мокин М.В., Отделенов В.А., Сычев В.А. Разработка онтологии технологических карт ведения пациентов многопрофильного стационара при моделировании медицинских технологических процессов // Искусственный интеллект и принятие решений. - 2014. - № 2. - С.68-77.
5. А.И. Молодченков Анализ результатов работы алгоритма автоматического построения условий, влияющих на выбор метода лечения // Искусственный интеллект и принятие решений. - 2014, № 4, - С. 52-58.

Прочие публикации

6. Малышевский А.А., Молодченков А.И., Хачумов В.М. Визуализация медицинских технологических процессов и многомерных данных для поддержки принятия решений // Информационные и телекоммуникационные системы: информационные технологии в научных и образовательных процессах. Материалы республиканской научно-практической конференции. Махачкала: ДНЦ РАН. 2009 г. - 268 с.
7. Молодченков А.И. Применение программных средств PsyExp для создания экспертных систем в области практической психологии. //Труды VIII международной конференции Интеллектуальный анализ информации ИАИ-2008, Киев: Просвіта, 2008. - С. 298-307.
8. Молодченков А.И. Экспертная система "Психология".//Труды одиннадцатой национальной конференции с международным участием КИИ-2008, Том 3, Дубна, ЛЕНАНД, 2008, С. 353-359
9. Молодченков А.И. Экспертные системы в области психодиагностики.//Труды XLIV Всероссийской конференции по проблемам математики, информатики, физики и химии, Москва, РУДН, 2008, С. 30-39.
10. И.А Тихомиров, А.И. Молодченков. Неоднородная семантическая сеть как средство представления знаний в предметной области "практическая психология". Теория и практика системного анализа: Труды I Всероссийской научной конференции молодых ученых. - Т. I. - Рыбинск: РГАТА имени П.А. Соловьева, 2010, стр. 45-49.

11. А.И. Молодченков. Применение AQ-алгоритма для персонификации лечебно-диагностических процессов. Теория и практика системного анализа: Труды I Всероссийской научной конференции молодых ученых. - Т. I. - Рыбинск: РГАТА имени П.А. Соловьева, 2010, стр. 79-84.
12. Молодченков А.И. Формализация описания лечебно-диагностических процессов. Нечеткие системы и мягкие вычисления: сб. ст. Третьей Всероссийской научной конференции: В 2 т. Т I , Волгоград, 2009, с. 103-109.

ЛИЧНЫЙ ВКЛАД АВТОРА В СОВМЕСТНЫЕ ПУБЛИКАЦИИ

В работе [2] предложено описание базы знаний для формализации знаний с помощью неоднородных семантических сетей. В [3] предложен метод синтеза плана лечения на основе прецедентов. В [4] предложена онтология медицинских технологических процессов на примере бронхиальной астмы. В [6] предложен способ формализации медицинских технологических процессов. В [10] предложены новые типы связей для формализации знаний в области психологии с помощью неоднородных семантических сетей.