



OSTIS-2016

(Open Semantic Technologies for Intelligent Systems)

УДК 004.89

КОНЦЕПТУАЛЬНОЕ ПРОЕКТИРОВАНИЕ УПРАВЛЯЕМОЙ СИСТЕМЫ ПОДДЕРЖКИ ИНТЕЛЛЕКТУАЛЬНОЙ ДЕЯТЕЛЬНОСТИ

Клещев А.С., Шалфеева Е.А.

*Институт автоматики и процессов управления,
г. Владивосток, Россия*

kleshev@iacp.dvo.ru

shalf@iacp.dvo.ru

Данная работа посвящена проблеме управления качеством баз знаний, используемых в процессе принятия решений повседневной интеллектуальной деятельности специалистов. Выделены Технологические этапы автоматизации интеллектуальной деятельности конкретной предметной области и предложена концептуальная архитектура системы автоматизации, которая позволяет решать множество взаимосвязанных интеллектуальных задач известных классов с использованием управляемых баз знаний.

Ключевые слова: автоматизация интеллектуальной деятельности; управление базами знаний; решатель задачи; концептуальный архитектурный проект.

Введение

Автоматизация профессиональной интеллектуальной деятельности и управления ее качеством требует не только решения проблемы поддержки интеллектуальной деятельности, но и решения проблемы правильности применения знаний при этом. Ключевыми аспектами современной парадигмы автоматизации интеллектуальной деятельности являются: понятные специалистам базы знаний (БЗ); полезные БЗ в течение всего времени их эксплуатации; возможность включения в них новых научных результатов. Недостаточное внимание поддержке процесса управления качеством баз знаний, используемых в процессе принятия решений, по-видимому, объясняет тот факт, что на сегодняшний день экспертные системы не «встроены» в организационную структуру учреждений такой, например, отрасли, как медицина, для систематической поддержки повседневной интеллектуальной деятельности специалистов [Жариков, 2008; РусБИТех, 2013].

Целью настоящего исследования является разработка концептуального архитектурного проекта системы автоматизации предметной области с интеллектуальной деятельностью (на примере медицинской области), в которой решается множество взаимосвязанных задач известных классов и требуется управление множеством взаимосвязанных баз знаний.

1. Место управления качеством в процессе автоматизации произвольной интеллектуальной деятельности

Поскольку полезность интеллектуальной деятельности определяется ее качеством, то параллельно с повседневным выполнением интеллектуальной деятельности, должно осуществляться управление ее качеством. Качество интеллектуальной деятельности определяется величиной риска ошибочных решений и характером ущерба от последствий таких ошибочных решений. Качество интеллектуальной деятельности в значительной степени зависит от качества используемых знаний.

Поэтому процесс управления качеством интеллектуальной деятельности состоит, прежде всего, в повышении качества знаний, используемых специалистами - совершенствования правильности знаний (при правильном применении знаний среди множества решений должно быть правильное), точности знаний (при правильном применении знаний получается единственное правильное решение), а также правильности и точности их применения.

Для оценки текущего уровня качества интеллектуальной деятельности и управления этим качеством обычно используются различного рода отчеты о результатах деятельности, включающие информацию о решениях, принимаемых специалистами в процессе своей деятельности, и *заклучения* (подтвержденные правильные решения

задач), такие отчеты могут быть использованы ответственными лицами (группой экспертов высокой квалификации), в частности, для выявления неверных или несовершенных знаний.

Поскольку для автоматизации интеллектуальной деятельности необходима поддержка не только повседневного выполнением интеллектуальной деятельности, но и управления ее качеством, требуется создавать комплекс из экспертных систем и средств обеспечения качества их решений. Сложность общей задачи управления качеством используемых в многоэтапном процессе принятия решений диктует необходимость разбиения интеллектуальной деятельности на отдельные задачи, с которыми связаны задачи управления качеством используемых знаний. (Для управляемости качеством знаний, используемых на каждом этапе принятия решения, целесообразно использовать принцип системного анализа, диктующий разбиение сложной задачи на ее составляющие - менее сложные.) Поэтому в составе автоматизирующих комплексов предпочтительно иметь множество экспертных систем (ЭС), предназначенных для решения задач разных классов.

В медицине классов таких задач несколько: диагностика, лечение, прогноз изменения состояния пациента, наблюдение за ним (мониторинг). По аналогии с другими предметными областями совокупно эти задачи составляют единую задачу управления сложной системой [Гаврилова и др., 2000; Джексон, 2001] (здесь - задачу управления здоровьем пациента).

2. Идентификация автоматизируемых интеллектуальных задач

Поскольку деятельность в предметной области включает в себя повседневную, в том числе интеллектуальную, контроль принимаемых решений и деятельность по управлению знаниями [Клещев и др., 2015], то (для обеспечения поддержки всех этих трех видов деятельности) при системном анализе выполняется:

- идентификация интеллектуальных задач в предметной области;
- разработка онтологий для каждой такой задачи;
- постановка каждой интеллектуальной задачи;
- разработка или выбор методов решения и соответствующих им алгоритмов для интеллектуальных задач.

По завершении этих работ целесообразно приступать к концептуальному проектированию системы автоматизации и к инженерии баз знаний, причем переход к инженерии баз знаний возможен после реализации ряда подсистем для инженерии знаний, выделенных при концептуальном проектировании [Клещев и др., 2015]. Для того

чтобы при проведении системного анализа для автоматизации профессиональной интеллектуальной деятельности рационально выделять отдельные подзадачи, методы решения которых уже известны, требуется обратиться к современной классификации постановок задач, охватывающей весь спектр задач, относящихся к области разработки систем, основанных на знаниях. Предложенная в [Клещев и др., 2015] многоуровневая классификация задач даёт возможность в процессе системного анализа переходить к математическим постановкам различных задач, конкретизируя те понятия и их абстракции, которые содержатся в этой классификации.

Среди множества подзадач задачи управления организмом в медицине *диагностика* в сочетании с *запросом дополнительной информации* и *составлением плана действий* с параллельно осуществляемым *прогнозом результата воздействий* наиболее очевидны.

В задаче *диагностики* дано: результаты наблюдения признаков, характеристик системы и произошедших событий (R_{Σ}); база знаний о клинической картине каждого заболевания, база знаний о нормальных значениях наблюдаемых признаков (KB_{Σ}).

Требуется найти: все возможные причинно-следственные модели системы, включающие диагноз D ($AS_{\Sigma}(R_{\Sigma})$) и согласованные с результатами наблюдений R_{Σ} , относительно которых все предложения из БЗ KB_{Σ} истинны.

В задаче *запроса дополнительной информации для распознавания* дано: база знаний о клинической картине каждого заболевания (KB_{Σ}), удовлетворяющая условию разделимости классов (диагнозов); R_{Σ} , такое, что для него множество гипотез ($H_{R,KB,\Sigma}$) имеет мощность не меньше двух.

Требуется найти: такой запрос дополнительной информации (Q_{Σ}) для результатов, что множество гипотез ($H_{R',KB,\Sigma}$), построенное для обновленного R_{Σ} , имеет меньшую мощность.

В задаче *лечения (планирования воздействий)* дано: значения характеристик системы (R_0); диагноз D ; желаемый результат наблюдений признаков - состояние *норма* (R^{cond}), база знаний о схемах лечения, база знаний о лечебных средствах - средствах медикаментозного воздействия).

Требуется найти: множество медикаментозных воздействий ($R_{ev} = \{m_1, \dots, m_k\}$), при которых новые результаты наблюдения признаков (как функции времени в соответствии с утверждениями базы знаний) будут удовлетворять условиям R^{cond} .

В задаче *прогноза результата воздействий* дано: значения характеристик системы (R_0); диагноз D , планируемые воздействия (R_{ev}); конечное множество моментов времени T ; база знаний о лечебных средствах.

Требуется найти: значения всех признаков во все моменты времени T в причинно-следственной модели $(AS_{\Sigma}(R_O, D, R_{ev}))$ [Клещев и др., 2015a].

Эти классы задач соответствуют подзадачам повседневной деятельности. Поскольку медицина относится к тем областям деятельности, где предусмотрена процедура подтверждения решений специалиста, для повышения уровня используемых знаний применим Метод монотонного усовершенствования БЗ [Клещев и др., 2015], называемый в данном исследовании *управлением знаниями*. Решение, принимаемое врачом, в дальнейшем признается правильным или ошибочным (например, правильным - по итогу выздоровления пациента в результате лечения, ошибочным - если выздоровление произошло только в результате применения другого, более действенного лечения).

Для монотонного усовершенствования ставшее известным правильное решение конкретной задачи (прецедент) сравнивается с объяснением, сформированным системой, и, если объяснение было неправильным, то пару <известные данные об объекте, заключение о правильности> следует использовать как прецедент для исправления базы знаний системы.

Критерий для оценки базы знаний по диагностике - наличие правильного (согласно заключению) диагноза в объяснении списка альтернативных диагнозов, составленном ЭС.

Минимальный критерий для оценки базы знаний о лечении - наличие каждого рекомендованного лечебного средства (ЛС) в правильно назначенном лечении (из прецедента) в списке R_{ev} альтернативных рекомендаций, составленном ЭС.

Схематическая постановка задачи модификации знаний (для задачи лечения) такова. Дано: Объяснение способа лечения, сформированное системой $(R_{ex} \cup R_O \cup D, R_{ev}, R'_{ex})$ и согласованное с базой знаний версии 1, прецедент $(R_{ex} \cup R_O \cup D, R^{prec}_{ev}, R^{prec}_{ex})$, согласно которому желаемое состояние было достигнуто с помощью воздействия, отсутствующего в предложенном системой списке возможных лечебных воздействий $(m_i \in R^{prec}_{ev} \text{ и } m_i \notin R_{ev})$.

Требуется найти: базу знаний версии 2, обеспечивающую Объяснение списка возможных лечебных воздействий, включающего воздействие m_i $(R_{ex} \cup R_O \cup D, R^{prec}_{ev}, R^{prec}_{ex})$ [Шалфеева и др., 2015].

Подобные классы задач соответствуют подзадачам регулярной деятельности управления качеством знаний, сопутствующей повседневной деятельности. Деятельность по управлению качеством знаний должна осуществляться экспертами, которым требуются удобные средства формирования обучающей выборки из

прецедентов, средства автоматического формирования на основе онтологии очередного варианта модификации БЗ и средства его оценивания.

Важно убедиться, что получен вариант БЗ, позволяющий давать правильные объяснения для всех задач, уже имеющихся на этот момент в базе прецедентов.

3. Концептуальное проектирование

Концептуальное проектирование системы состоит из множества шагов [Pressman, 2001].

В медицине на шаге идентификации (сколько и каких потребуется) программных подсистем для автоматизации повседневной «обычной» деятельности должны быть предусмотрены автоматизированные «рабочие места» (АРМ) регистратора, лаборанта, позволяющие добавлять значения постоянных или измеренных значений признаков в ИБ как часть подсистемы документирования.

На шаге идентификации подсистем для автоматизации повседневной интеллектуальной деятельности (в отдельном медицинском звене или целой отрасли) должны быть идентифицированы все базы знаний, используемые для поддержки решения каждой их интеллектуальных подзадач, должны быть предусмотрены «решатели», генерирующие на основе знаний:

- предложения по получению наиболее показательного признака,
- объяснения диагноза, обусловленного значениями известных наблюдений,
- объяснения выбора одних средств медикаментозного воздействия и отклонения других,
- прогноз развития значений наблюдаемых признаков для дальнейшего их мониторинга.

На шаге идентификации программных подсистем для инженерии баз знаний должны быть идентифицированы редакторы всех баз знаний (о клинической картине каждого заболевания, о нормальных значениях наблюдаемых признаков, о схемах лечения, о ЛС и их воздействиях). Все редакторы обеспечивают работу с информацией в терминах зафиксированной онтологии предметной области. Инженерия баз знаний требует включать в комплекс также и «рабочее место» эксперта для поддержки анализа проверенных решений задач врача и отбора прецедентов, подсистему для проверки соответствия очередного варианта базы знаний - базе прецедентов, подсистемы индуктивного формирования знаний (в частности, знаний о клинической картине каждого заболевания), подсистемы проверки соответствия знаний базе прецедентов.

На шаге построения схемы взаимосвязи всех подсистем в единой системе планируются

интерфейсы взаимодействия всех идентифицированных подсистем с соответствующими типами пользователей, с информационными ресурсами (базами знаний, текущими историями болезней, базами прецедентов и хранилищами другой вышеупомянутой информации), друг с другом в некоторых случаях [Клещев и др., 2015].

В результате концептуального проектирования системы автоматизации появляется модель системы (top-level architecture flow diagram), состоящей из подсистем, связанных информационными потоками друг с другом, с хранилищами информации и с внешним окружением [Pressman, 2001]. Пример фрагмента такой модели для автоматизации медицинской деятельности представлен на рис. 1.

На рисунке показано, с какими частями системы автоматизации взаимодействует специалист, решающий задачи, с какими – команда управления качеством БЗ (эксперты, ответственные за качество знаний). Подсистема-редактор доступна не только врачам,

но и медперсоналу, который с решателями не работает, а занимается только вводом и редактированием данных о пациенте. За кадром осталась подсистема документооборота со специализированными редакторами баз данных и компонентами, генерирующими нужные отчеты по хранимым данным в этих базах для «команды управления» деятельностью (руководителей разных звеньев).

Специалистам каждого первичного звена предоставляется возможность взаимодействовать с совокупностью решателей ЭС, устанавливаемой на их рабочих местах, вводить информацию с помощью специализированных редакторов.

Решатель диагностики формирует один (или несколько возможных диагнозов) и их объяснения. Врач может не согласиться ни с одним из них и поставить свой диагноз. Решатель лечения должен строить план лечения для каждого из сгенерированных возможных диагнозов и их объяснения.

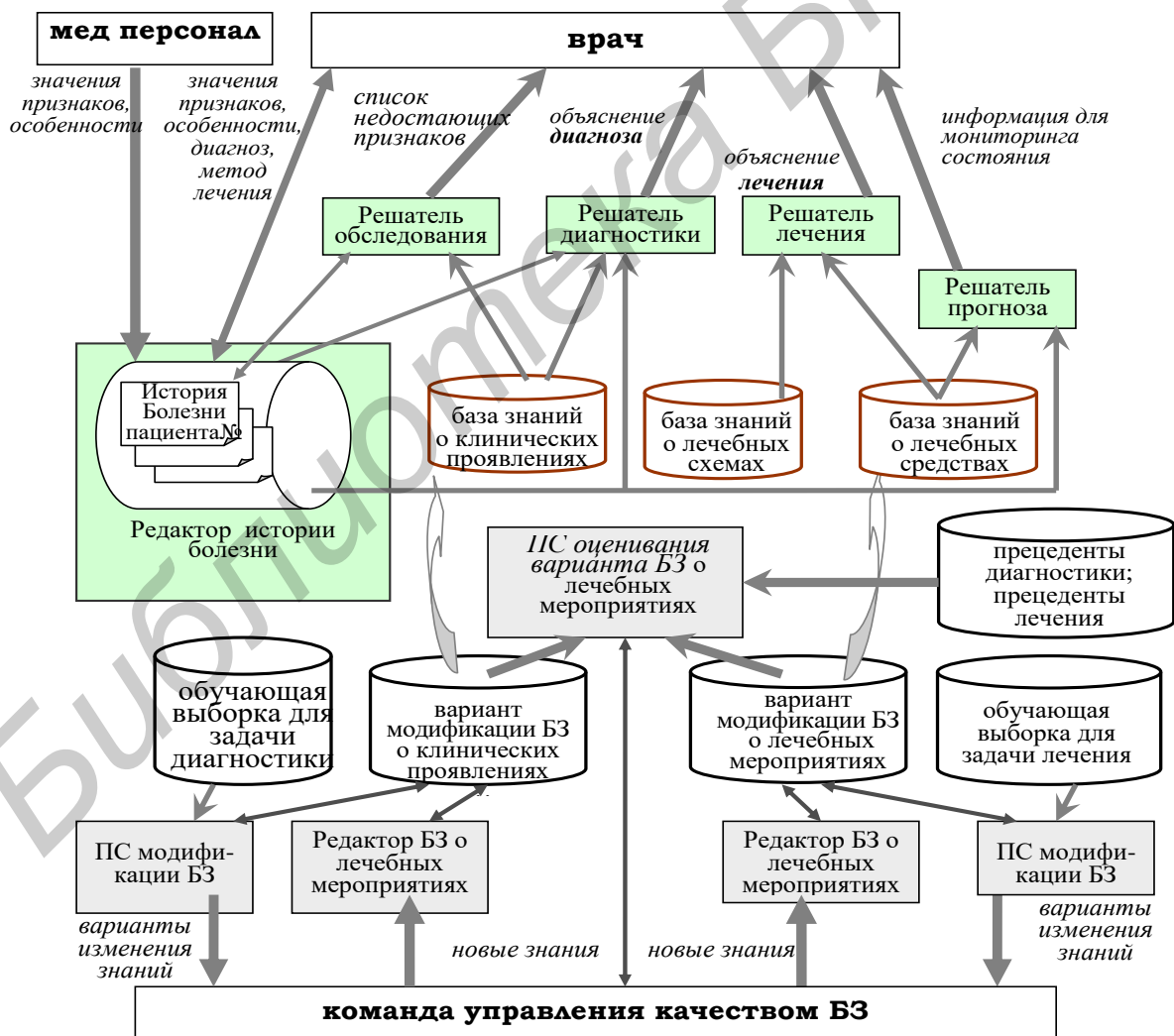


Рисунок 1 – Фрагмент концептуальной архитектуры системы автоматизации медицинской повседневной деятельности

Первичной разработкой баз знаний занимаются эксперты. Они же занимаются их сопровождением и управлением, являясь пользователями подсистем управления качеством знаний в процессе всего периода эксплуатации системы. В медицине с каждым профилем (офтальмология, урология, и т.д.) связана своя команда управления качеством баз знаний. Работая с однотипными АРМами, эксперт каждого профиля «отвечает» за качество своей совокупности баз знаний. Качество улучшаемых БЗ будет зависеть от сложности задач, которые специалистам приходится решать. Привлечение для управления знаниями дополнительных специалистов экономически оправдано в случае облачной реализации систем автоматизации (для уровня всей отрасли), если на центральном сервере установлены решатели, базы знаний, архив решенных задач и все подсистемы для управления.

Предложенная концептуальная архитектура базируется на «модели репозитория» и соответствует взаимозависимости автоматизируемых процессов в рассматриваемой предметной области. Всем пользователям интеллектуальной программно-информационной системы (врачам, медперсоналу, экспертам по знаниям и экспертам по подтверждению принятых решений) обеспечивается доступ к программным средствам.

4. Подход к реализации комплекса для поддержки принятия решений с управляемой базой знаний

Требования, относящиеся к способам формирования и использования баз знаний и другой хранимой информации, к доступности компонентам системы диктуют выбор «среды разработки» для системы автоматизации медицинской деятельности. Одно из главных требований относится к возможности представления знаний о клинической картине заболеваний не только с учетом некоторой динамики развития признаков, но и с учетом вариантов развития такой динамики. Иерархические однородные семантические сети отвечают предъявляемым требованиям к естественной для специалистов структуре описания знаний о клинических проявлениях заболеваний со множеством вариантов развития динамики признаков (рис. 2).

Другое требование связано с возможностью отслеживания прецедентов и своевременного их применения для усовершенствования знаний. В медицине важно использовать информацию о новом прецеденте, обнаруженном в любом учреждении отрасли, чтобы улучшенной с его помощью базой знаний могли пользоваться все. Поэтому облачный доступ к средствам оценивания и модификации БЗ, является наиболее адекватным выбором.

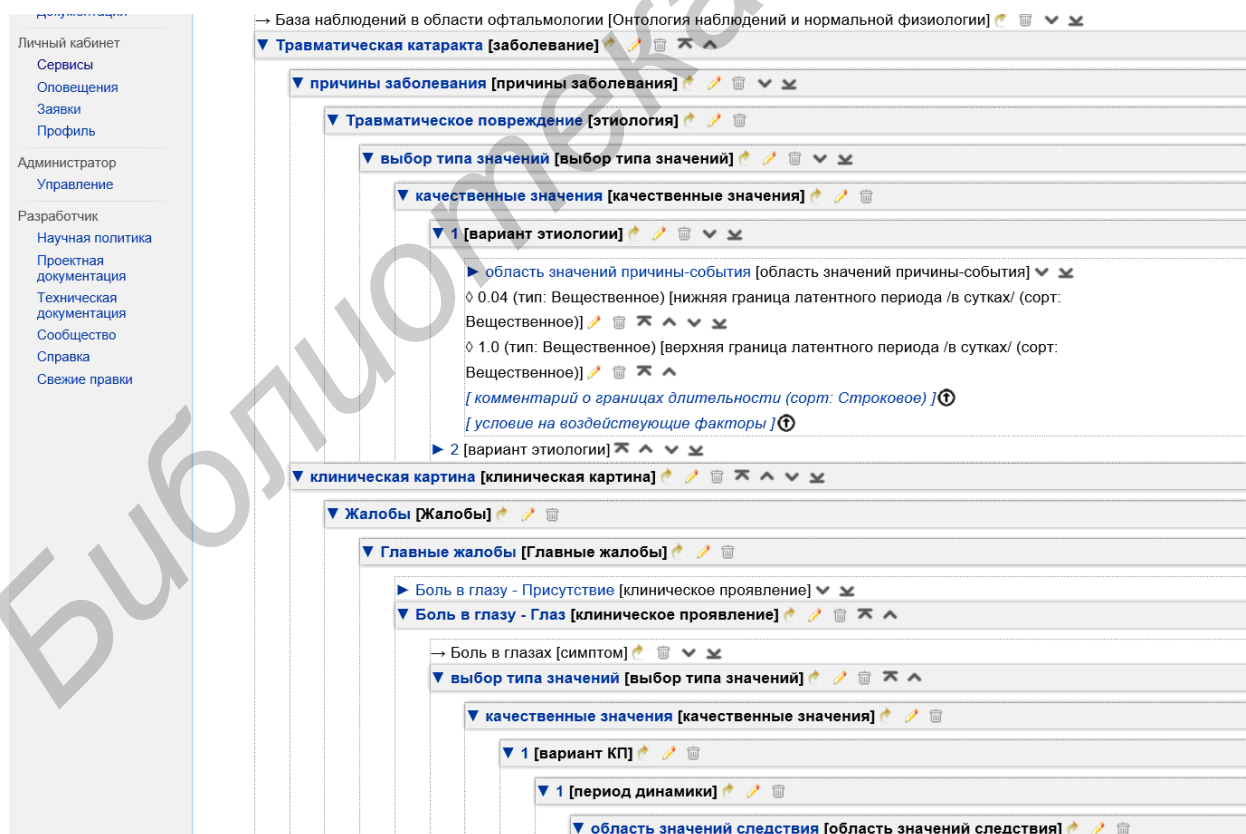


Рисунок 2 – Фрагмент описания клинической картины заболевания с учетом вариантов развития динамики признаков

Примером облачной платформы для реализации программ, обрабатывающих информацию в виде иерархических однородных семантических сетей, поддерживающих естественную для специалистов структуру информации и терминологии, является IASPaas [Грибова и др., 2011]. Технология IASPaas подразумевает создание специализированных решателей, обрабатывающих знания, представленные декларативно (в виде иерархических семантических сетей). Технология IASPaas поддерживает декларативность представления программных компонентов для создаваемых ИС. При разработке автоматизирующего медицинского комплекса каждый решатель реализуется чаще как самостоятельный облачный сервис, формирующий объяснение гипотез о решении задачи. Структура подаваемых на вход входных инфоресурсов, в том числе баз знаний, описывается в этой декларации. Замена версий подаваемой на вход базы знаний позволяет оперативно приспособлять решатель к уточняемым знаниям предметной области. Фрагмент примера входного информационного ресурса, соответствующего указанной в декларации метаинформации, и обрабатываемого этим решателем, показан на рис. 3. Пример декларации решателя задачи лечения, сформированной в среде IASPaas, показан на рис. 4.

При декларировании решателя указываются форматы обрабатываемой ими информации (входной и выходной). Входными для задачи

лечения будут истории болезни и используемые знания, а выходные – объяснение. Объяснение, в котором действительно заинтересован специалист в реальной практике, не должно давать свой вариант назначения ЛС пациенту. Врач нуждается в информации о том, какие конкретные ЛС не следует назначать данному пациенту с его «особенностями», потому что анализ противопоказаний десятков, а то и сотен известных ЛС с нужным фармакологическим воздействием связан с существенными затратами времени. Онтология объяснения решателя лечения формируется, исходя из этих обстоятельств.

Если решение задачи терапевтического лечения к получению списка ЛС, рекомендуемых пациенту с указанным диагнозом, к последующему устранению из него ЛС, не подходящих конкретному пациенту (с указанными возрастом и прочими «особенностями» организма); далее (особенно в случае сочетанной патологии) важно провести анализ множества рекомендованных ЛС с учетом их сочетаемости при одновременном приеме. Пример результата, сформированного решателем IASPaas, и сохраняемого для использования системой управления знаниями, показан на рис. 5.

Пример результата, сформированного подсистемой проверки соответствия очередной версии базы знаний - базе прецедентов, показан на рис. 6.

The screenshot displays a web-based interface for a knowledge base. It features two main sections, each for a different drug:

- Ампициллин [ЛС]:**
 - Ампициллин (тип: Строковое) [названиеЛС (сорт: Строковое)]
 - противопоказания [противопоказания]:**
 - бронхиальная астма (тип: Строковое) [бронхиальная астма (тип: Строковое)]
 - лейкопения (тип: Строковое) [лейкопения (тип: Строковое)]
 - почечная недостаточность (тип: Строковое) [почечная недостаточность (тип: Строковое)]
 - поллиноз (тип: Строковое) [поллиноз (тип: Строковое)]
 - 0.1 (тип: Вещественное) [нижний возраст (сорт: Вещественное)]
 - Пенициллины [фармакологическая группа]
 - Взаимодействие с другими ЛС [Взаимодействие с другими ЛС]
 - показания [показания]
 - Фармакологическое действие [Фармакологическое действие]
 - торговые названия [торговые названия]
- Бенциклан [ЛС]:**
 - Бенциклан (тип: Строковое) [названиеЛС (сорт: Строковое)]
 - показания [показания]
 - противопоказания [противопоказания]:**
 - Гиперчувствительность (тип: Строковое) [Гиперчувствительность (тип: Строковое)]
 - почечная недостаточность (тип: Строковое) [почечная недостаточность (тип: Строковое)]
 - дыхательная недостаточность (тип: Строковое) [дыхательная недостаточность (тип: Строковое)]

Рисунок 3 – Пример базы знаний о воздействиях ЛС

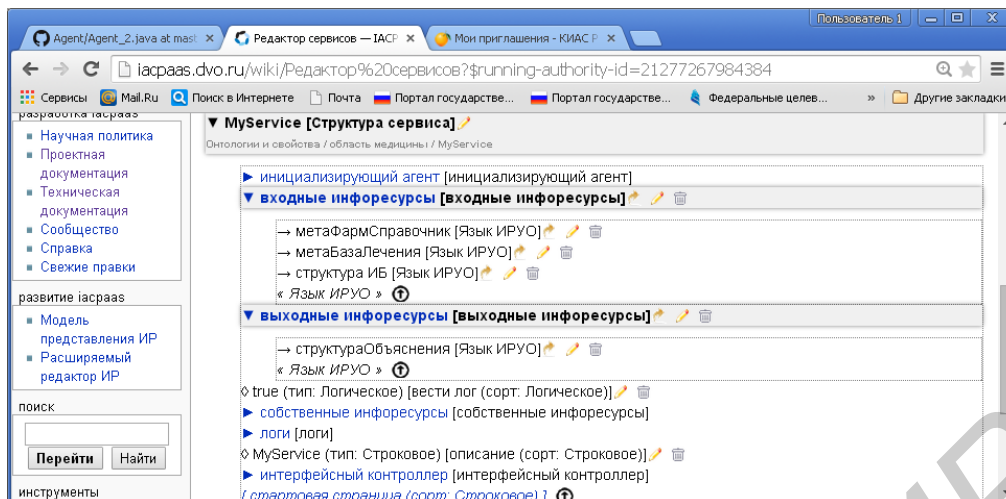


Рисунок 4 – Пример декларации решателя, реализующего решатель лечения

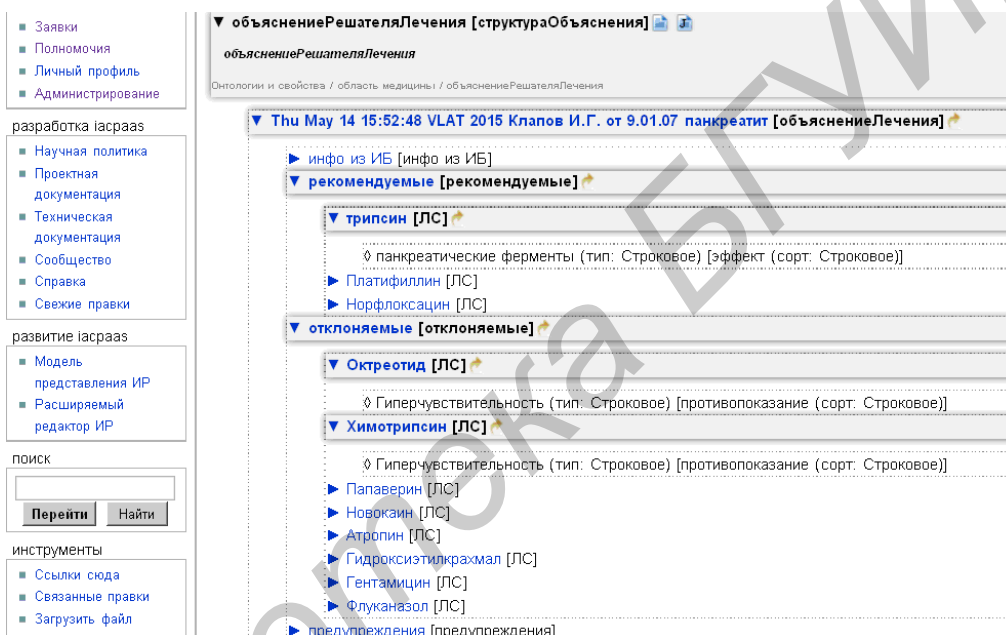


Рисунок 5 – Фрагмент объяснения назначения лечебных средств пациенту

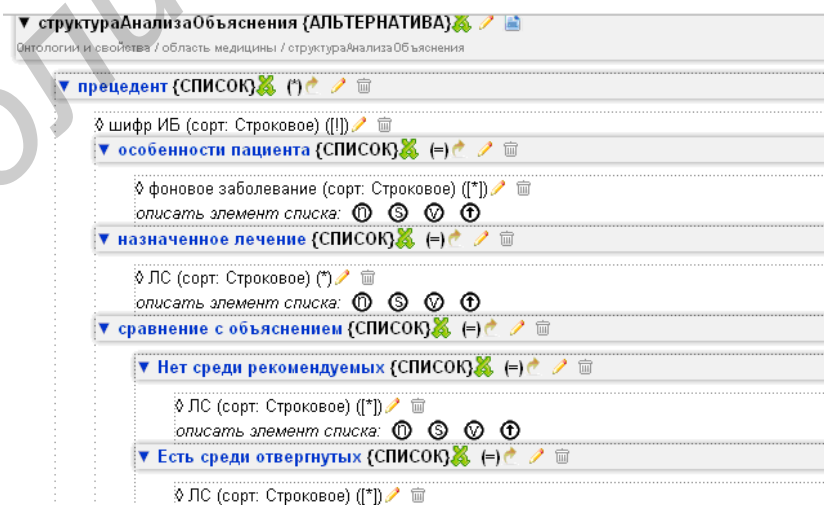


Рисунок 6 – Структура отчета о сравнении объяснения и правильного назначения

CONCEPTUAL DESIGN OF MAINTAINABLE SYSTEM OF INTELLIGENT ACTIVITY AUTOMATION

Kleschev A. , Shalfееva E.

*The Institute of Automation and Control
Processes, Vladivostok, Russia*

kleschev@iacp.dvo.ru

shalf@iacp.dvo.ru

This article is devoted to problem of quality control of knowledge bases used in decision-making process in daily intellectual activity. Technological stages of intellectual activity automation for some domain are allocated. The conceptual architecture of software system allowing to solve the set of interconnected problems of known classes using knowledge bases is proposed.

Introduction

The purpose of this research is development of the conceptual architectural design software of system of automation of medical activity - domain in which the set of the interconnected problems of known classes is solved and knowledge bases control is required.

Main Part

In the main part of the article the «place» of quality management in any intellectual activity automation is presented; identification of tasks in the domain is executed. The top-level architecture flow diagram (for medical activity automation) consisting of the subsystems connected by information streams with each other with storages of information and with an external environment is constructed.

The approach to implementation of a software of system of automation of medical activity taking into account requirements to ways of formation and use of knowledge bases and availability of all components of system for maintenance is offered.

Conclusion

The technological stages of automation of intellectual activity in concrete domain can be generalized to any area taking into account the principle of the system analysis and a modern paradigm of automation. The offered conceptual architecture of system of automation allows to solve a set of the interconnected intellectual problems of known classes where knowledge bases control is required. Feasibility of the offered approach is tested within the basic IACPaaS technology for problems of treatment of health for several medical specialization areas.

Заключение

Технологические этапы автоматизации интеллектуальной деятельности конкретной предметной области (медицинского первичного звена или отрасли) могут быть обобщены до произвольной области, поскольку выделены с учетом принципа системного анализа и современной парадигмы автоматизации. Предложенная концептуальная архитектура системы автоматизации позволяет решать при ее внедрении множество взаимосвязанных интеллектуальных задач известных классов, где требуется управление совокупностью баз знаний. Реализуемость предлагаемого подхода опробована в рамках базовой технологии IACPaaS для задач управления здоровьем по нескольким профилям (в частности, офтальмологические заболевания, заболевания ЖКТ).

Работа выполнена при частичной финансовой поддержке грантов РФФИ № 15-07-03193: «Облачные технологии обеспечения жизнеспособности и качества интеллектуальных систем» и № 14-07-00270 "Медицинские облачные сервисы"

Библиографический список

- [Гаврилова и др., 2000] Гаврилова, Т. А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. СПб: Питер, 2000, 384 с.
- [Джексон, 2001] Джексон П. Введение в экспертные системы: Уч. пос. – М.: Издательский дом "Вильямс", - 2001. – 624 с.
- [Грибова и др., 2011] Грибова В.В., Клещев А.С., Крылов Д.А. и др. Проект IACPaaS. Комплекс для интеллектуальных систем на основе облачных вычислений. Искусственный интеллект и принятие решений. 2011. №1. С.27-35.
- [Жариков, 2008] Жариков О.Г., Литвин А.А., Ковалёв В.А. Экспертные системы в медицине. // Медицинские новости. 2008. №10. С. 15-18.
- [Клещев и др., 2015] Клещев А.С., М.Ю. Черняховская, Шалфеева Е.А. Особенности автоматизации интеллектуальной деятельности Научно-техническая информация. Сер.2., 2015, No. 1, с. 10–20.
- [Клещев и др., 2015a] Клещев А.С., Шалфеева Е.А. Онтология задач интеллектуальной деятельности // Онтология проектирования. Самара: "Новая техника", 2015. № 2 (16). с. 179-205
- [РусБИТех, 2013] НПО РусБИТех. Единое информационное пространство в здравоохранении РФ. 2011. <http://www.myshared.ru/slide/96415/> (Актуально на 30.11.2015).
- [Шалфеева и др., 2015] Шалфеева Е.А., Бочкарёв А.С. Концептуальная архитектура и подход к реализации медицинского сервиса, использующего управляемую базу знаний // Материалы Всероссийской конференции с международным участием "Знания – Онтологии – Теории" (ЗОНТ-2015). 6-8 октября 2015 г., Новосибирск, Институт математики им. С.Л. Соболева СО РАН. - Новосибирск: ООО «Технотрейд», 2015. Том 2. – С. 196-205.
- [Pressman, 2001] Pressman R.S. Software Engineering: Practitioner's Approach. Fifth edition. - McGraw-Hill Inc., 2001. - 860 p.