



щелчками реакций реагирования на слова, что открывает широкие перспективы в практике лечения нервно-психических заболеваний, организации эффективной работы человека-оператора и т.д.; при преимущественном раздражении слухового анализатора активизируются центр моторной речи, а также другие отделы коры.

ЛИТЕРАТУРА

1. *Воронов А.В., Горбач Т.А.* Исследование когнитивности человека. Методика и электроэнцефалографические исследования // Информатика и системы управления. – 2007. – №1. – С.45-56.
2. *Равич-Щербо И.В. и др.* Психогенетика. Учебник / И.В. Равич-Щербо, Т.М. Марютина, Е.Л. Григоренко; под ред. И.В. Равич-Щербо. – М.: Аспект-Пресс, 2000.
3. Атлас “Нервная система человека. Строение и нарушения”/ под ред. В.М. Астапова и Ю.В. Микадзе. – 5-е изд., перераб. и доп. – М.: ПЕР СЭ, 2006.

Статья представлена к публикации членом редколлегии Е.А. Нурминским.

УДК 004.891.3:61

© 2007 г. **Ф.М. Москаленко**

(Институт автоматизации и процессов управления ДВО РАН, Владивосток)

ПРОЕКТ КОМПЬЮТЕРНОГО БАНКА ЗНАНИЙ ПО МЕДИЦИНСКОЙ ДИАГНОСТИКЕ¹

Описано информационное и программное наполнение банка знаний по медицинской диагностике. Определены классы его пользователей и задачи, ими решаемые. Информационное наполнение содержит базы наблюдений, базы знаний о заболеваниях, и базы данных (пациентов). Программное наполнение содержит редакторы информации, а также решатель задачи медицинской диагностики.

Введение

Одним из приложений систем искусственного интеллекта являются системы медицинской диагностики. Их применение помогает врачу повысить качество своей работы. Краткий анализ разрабатываемых в последние

¹ Работа выполнена при финансовой поддержке ДВО РАН в рамках программы № 14 Президиума РАН, проект 06-И-П14-052, и в рамках проекта РФФИ № 06-07 89071 «Исследование возможностей коллективного управления в семантическом вебе информационными ресурсами различных уровней общности».

годы систем приведен в [1]. Он показал, что данные используемые в системах медицинской диагностики, основанные на знаниях онтологии предметной области, являются сравнительно простыми и одновременно не отражают такие повсеместно используемые врачами в их практике знания предметной области: о причинах заболеваний; о различных типах причинных связей между признаками и заболеваниями; о воздействиях событий на значения признаков при заболеваниях и у здоровых пациентов; о различных вариантах изменения значений признаков, зависящих от анатомо-физиологических особенностей пациентов.

Помимо того, одно из отрицательных свойств некоторых разрабатываемых систем – то, что круг их применения достаточно узок. Это обусловлено тем, что они представляют собой либо макетные версии, выполненные для каких-либо исследовательских целей, либо разработаны для определенного медицинского учреждения и недоступны за рамками его локальной сети. С другой стороны, системы медицинской диагностики, предоставляющие широкий доступ к своим ресурсам с применением современных сетевых технологий (Интернет), – например, DXplain [2] и «Диагностика преэклампсии» [3], не позволяют экспертам расширять используемые в них базы знаний.

Таким образом, актуальной задачей является разработка системы медицинской диагностики, основанной на знаниях экспертов и модели онтологии, учитывающей все приведенные выше особенности медицинских знаний, в которой их модель имеет форму, наиболее близкую к представлениям экспертов, и позволяет не только определять диагноз пациента, но и объяснять его. Такая система должна проводить диагностику за приемлемое для врача время (несмотря на то, что в ее основе лежит нетривиальная онтология медицинских знаний). Кроме того, такая система должна предоставлять доступ как можно большему числу пользователей как для проведения медицинской диагностики, так и для участия в накоплении и совершенствовании медицинских знаний о различных заболеваниях.

Цель данной работы – описание концепции сетевого ресурса по медицинской диагностике, обладающего указанными выше свойствами.

Теоретические предпосылки и общие принципы создания банка знаний по медицинской диагностике

За последние годы специалистами в области искусственного интеллекта и экспертами в области медицинской диагностики было разработано несколько моделей онтологии медицинской диагностики. Для объединения преимуществ этих моделей и создания модели онтологии, наиболее приближенной к представлениям знаний в области медицинской диагностики, были созданы и описаны онтология и ее модель [4, 5]. Данная онтология медицинской диагностики описывает острые заболевания, в ней учитывается взаимодействие причинно-следственных отношений различных

типов. Она близка к реальным представлениям медицины в Российской Федерации и описывает сочетанную и осложненную патологии, динамику патологических процессов во времени, а также воздействие лечебных мероприятий и других событий на проявления заболеваний. Моделью этой онтологии является небогатая система логических соотношений с параметрами, которая состоит из определений терминов модели действительности (неизвестных), определений терминов модели знаний (параметров), ограничений целостности неизвестных и параметров, а также соотношений между ними.

На основе данной модели онтологии в работе [1] поставлена общая и частная (при наложении на модель некоторых ограничений) задачи медицинской диагностики, а также алгоритм решения частной задачи. В работе [6] приводится распараллеленный алгоритм решения этой задачи, а также описана система, выполняющая оптимизирующее преобразование базы знаний о заболеваниях, в результате чего возможно сокращение количества гипотез о диагнозе. Результаты экспериментального исследования временной сложности оптимизированного алгоритма решения частной задачи медицинской диагностики, описанного в [6], приведены в работе [7].

Описанная онтология и основанный на ней алгоритм медицинской диагностики могут быть использованы при создании системы для обеспечения процесса согласованного решения комплекса задач по сбору, формализации, переводу в машиночитаемое представление, инженерии, хранению, управлению и обработке данных и знаний в области медицинской диагностики, это объединение всей информации в единый ресурс с возможностью удаленного доступа к нему многих пользователей. Разрабатываемый ресурс назовем *банком знаний по медицинской диагностике*. Его архитектура включает следующие компоненты:

информационное наполнение (далее – ИН), хранящееся на сервере банка знаний и содержащее знания (о наблюдениях и заболеваниях) и данные (истории болезней пациентов), сформированные согласно схемам знаний и данных, задаваемых применяемой моделью онтологии предметной области;

программное наполнение (далее – ПН), ориентированное на интеллектуальную поддержку пользователей банка, разработанное в соответствии с технологией «клиент-сервер» и включающее:

средства по редактированию данных из ИН (баз знаний наблюдениях и заболеваниях и базу историй болезни), выполняющие также роль подсистемы доверия, так как они обеспечивают просмотр накопленных знаний, применяемых при диагностике;

средства по обработке данных из ИН, – в частности, оптимизация базы знаний о заболеваниях, интеллектуальный решатель задачи медицинской диагностики;

административную подсистему.

Кроме того, отметим следующие особенности архитектуры:

ИН банка состоит из разделов, каждый из которых содержит знания и данные конкретной области медицинской диагностики (офтальмология, иммунология и т.п.);

поскольку алгоритм, описанный в [6], предназначен для реализации на многопроцессорном вычислительном комплексе, то часть средств ПН и все компоненты ИН находятся на сервере банка знаний, а подсистема, непосредственно выполняющая диагностику, – на кластере.

Пользователей разрабатываемого банка знаний можно разделить на следующие группы:

эксперты, которые формируют базы знаний о наблюдениях и о заболеваниях;

врачи, формирующие истории болезней пациентов и пользующиеся экспертной системой медицинской диагностики;

гости банка, которые могут знакомиться с применяемыми базами знаний, формируемыми экспертами;

администратор, выполняющий функции главного эксперта банка (на основе персональных баз знаний экспертов формирует общую базу знаний о заболеваниях для каждого раздела банка) и осуществляющий управление полномочиями пользователей.

Построение экспертной системы в соответствии с технологий «клиент-сервер» позволяет:

организовать коллективное сопровождение общих баз знаний (под управлением администратора) и ведение персональных баз знаний экспертов;

организовать коллективное накопление баз данных (историй болезни) для совершенствования баз знаний, научных и иных целей;

удаленно использовать единый централизованный банк знаний и блок логического вывода всеми заинтересованными пользователями (под контролем администратора);

разместить компоненты приложения там, где они будут функционировать наиболее эффективно;

вести централизованное сопровождение программного обеспечения банка знаний.

Из литературы [8, 9] уже известны попытки разработки распределенных медицинских приложений на основе Интернета. В работе [8] предложена схема клиент-серверной экспертной системы для сетей Internet/Intranet, которая может быть использована при решении задач телемедицины. Однако работа не содержит описания средств и методов администрирования системы, несмотря на то, что такой компонент является важной составляющей многопользовательских систем. Понятие «банк знаний» применяется узко – только к хранимой информации (набор баз знаний), тогда как в предлагаемом здесь подходе оно включает также и методы работы со

знаниями – методы их редактирования и интеллектуальные решатели задач. В работе не рассматриваются подсистемы доверия и объяснения результатов, также в ней отсутствует описание средств контроля за накоплением знаний. Хранение какой бы то ни было информации на стороне клиента (базы данных) вызывает сомнение, поскольку в этом случае пользователю приходится работать с системой только с одного и того же рабочего места (клиент перестает быть «тонким»). Можно отметить лишь то ее преимущество, что она ориентирована на применение широко распространенного формата XML для передачи данных между клиентом и сервером.

В работе [9] описывается разработка инструментария для компьютеризированной системы поддержки процесса диагностики, основанной на технологии *case-based reasoning*, а также помогающей в образовании, проведении виртуальных медицинских конференций и в управлении записями пациентов с иллюстративными историями болезней. Однако это медицинское мультимедийное информационное хранилище данных не предполагает наличия экспертной системы диагностики в качестве подсистемы, так как авторы планируют включать в систему лишь средства, позволяющие находить истории болезней пациентов с известными диагнозами, необходимые врачу для самостоятельной постановки диагноза. Предлагаемая же здесь концепция банка знаний предполагает, что содержимое ПН позволяет проводить интеллектуальную обработку знаний и данных.

Информационное наполнение банка знаний по медицинской диагностике

Как уже отмечалось, в каждом разделе ИН банка хранятся базы трех типов: база знаний о наблюдениях; базы знаний о заболеваниях (одна общая и, возможно, несколько индивидуальных баз разных экспертов); база историй болезней.

База знаний о наблюдениях задается экспертами на начальном этапе формирования знаний в банке, после чего ее содержимое используется при формировании базы знаний о заболеваниях и базы историй болезней пациентов. Однако эта база может дополняться в процессе эксплуатации банка новыми понятиями. В каждом разделе банка знаний присутствует единственная база знаний о наблюдениях, применяемая всеми пользователями. Ее структура задается используемой моделью онтологии [4,5] и может быть представлена в виде схемы (рис. 1).

На рис. 1 и последующих рисунках прямоугольниками обозначены понятия модели онтологии. Они связаны стрелками, указывающими направление зависимости – от понятия-предка к понятию-потомку. На рисунке выделены места слияния стрелок (кружки) и места их разделения (квадраты). Разделение стрелки, выходящей из понятия-предка, говорит о том, что этому понятию сопоставляется только одно из понятий-потомков. Перечеркнутое штрихом начало стрелки говорит о том, что понятию-

предку не обязательно сопоставляется понятие-потомок. Если некоторое понятие (понятие-потомок) обозначено несколькими прямоугольниками, расположенными друг над другом, – значит, соответствующему ему понятию-предку может быть сопоставлено несколько потомков, являющихся экземплярами рассматриваемого понятия. Пунктиром «БН» выделена часть понятий модели и их связей, используемых в следующей схеме.

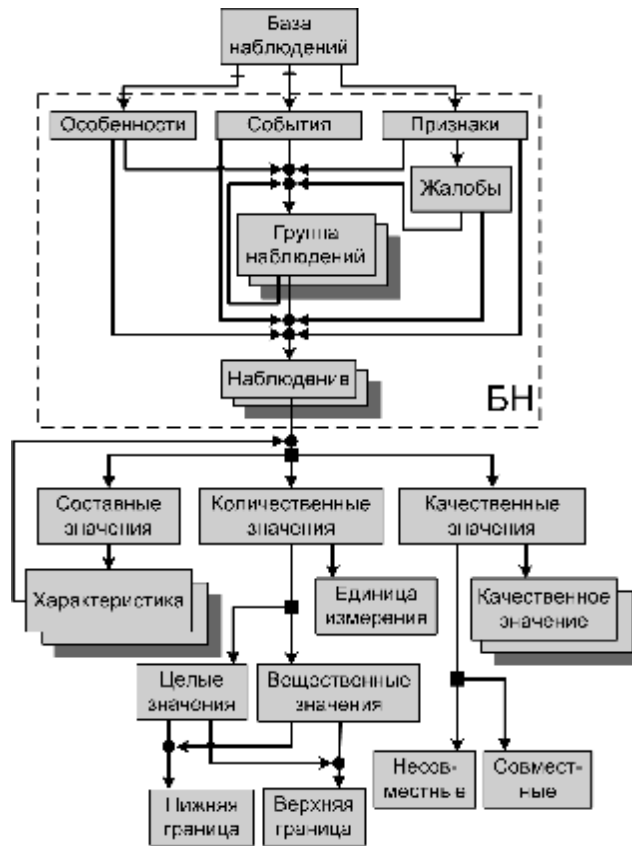


Рис. 1. Схема модели онтологии базы наблюдений в области медицинской диагностики.

Согласно описанной схеме базы наблюдений в саму базу заносятся названия признаков, особенностей, событий, наблюдений и их групп, названия характеристик, значения наблюдений.

База историй болезней формируется в соответствии с задаваемой моделью онтологии схемой, представленной на рис. 2. На рисунке пунктирным прямоугольником «БН» обозначена часть архитектуры модели онтологии, взятая с рис. 1. Отметим, что модель онтологии базы историй болезней пациентов строится на основе модели онтологии базы наблюдений, а сама база историй болезней формируется на основании базы наблюдений. История болезни представляет собой описание состояния пациента в моменты проведения осмотров и до поступления в клинику: задаются значения событий, признаков, анатомо-физиологических особенностей. События могут происходить с пациентом в различные моменты времени и иметь различные значения. Признаки в различные моменты времени также могут

иметь различные значения, а анатомо-физиологические особенности пациента имеют постоянные во времени значения.

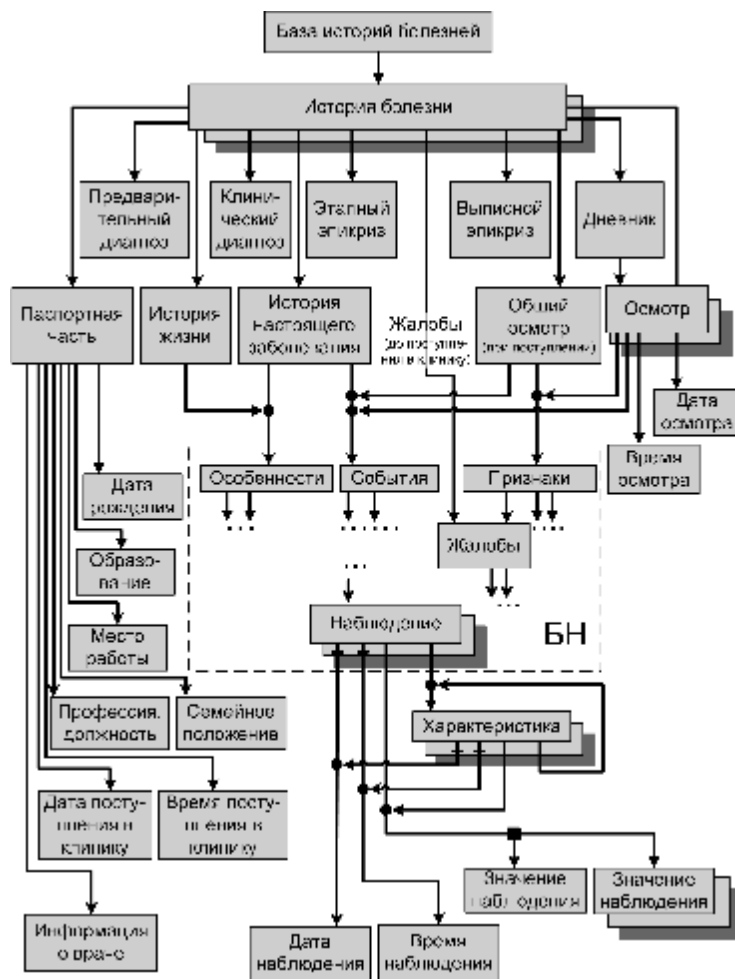
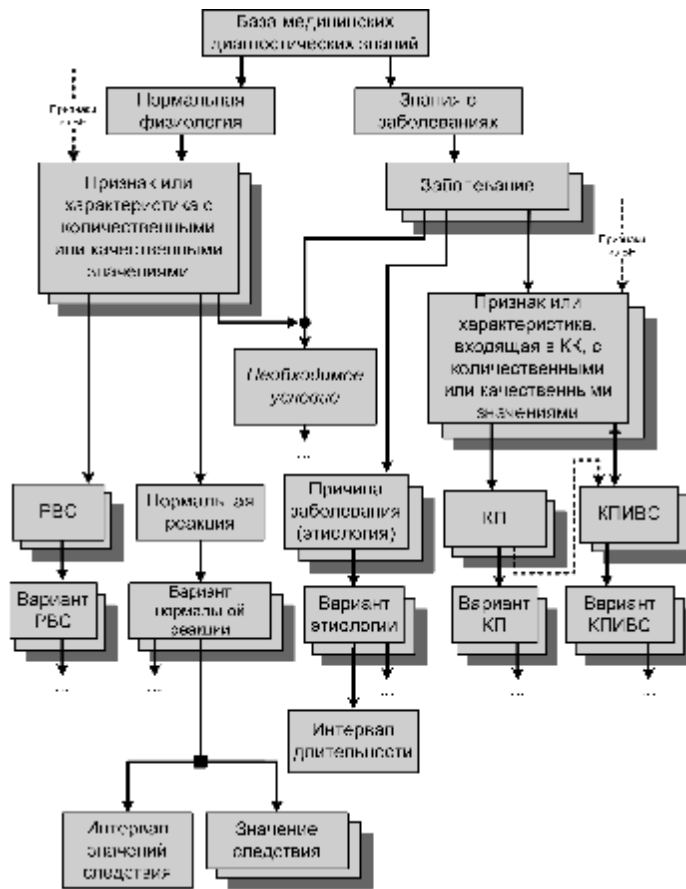


Рис. 2. Схема модели онтологии базы историй болезней пациентов.

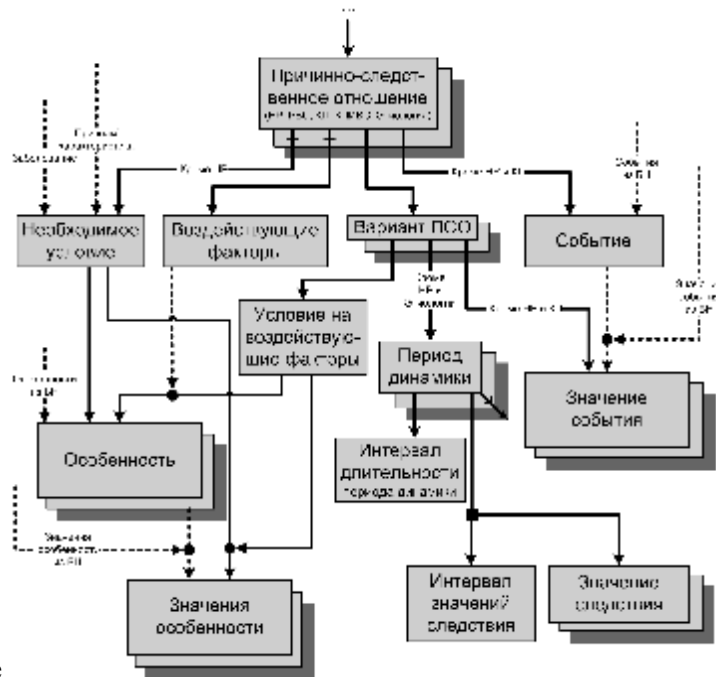
В каждом разделе банка знаний присутствует единственная общая база историй болезней, применяемая всеми пользователями.

База знаний о заболеваниях содержит описания заболеваний (значения признаков при заболеваниях задаются клиническими проявлениями и клиническими проявлениями, измененными воздействием событий), их причин (задаются этиологиями) и нормального состояния пациента (описываются нормальные реакции и реакции на воздействие событий).

Схема базы знаний о заболеваниях определяется моделью онтологии и представлена на рис. 3(а,б). На рис. 3а отображена первая часть схемы базы заболеваний. Понятие *Необходимое условие* расшифровано на рис. 3б. Кроме того, на этом рисунке продолжено описание понятий *РВС*, *Нормальная реакция*, *КП*, *КПИВС* и *Этиология*, сгруппированных в понятие *Причинно-следственное отношение*, а также понятий *Вариант РВС*, *Вариант нормальной реакции*, *Вариант КП*, *Вариант КПИВС* и *Вариант Этиологии*, сгруппированных в понятие *Вариант ПСО*.



а) начало



б) окончание

Рис. 3. Схема модели онтологии базы диагностических знаний.

Пунктирными стрелками с подписями, входящими в некоторые понятия на схеме модели онтологии знаний о заболеваниях, указывается, что при заполнении базы знаний о заболеваниях значения, подставляемые в эти понятия, берутся из сформированной базы наблюдений.

Отметим, что для каждой совокупности взаимосвязанных заболеваний (раздела) в ИН присутствует несколько баз знаний о заболеваниях: одна – общая и несколько (возможно, ни одной) индивидуальных. Каждый эксперт снабжается индивидуальной базой знаний о заболеваниях в разделе ИН, с которым он работает. Впоследствии индивидуальные базы знаний объединяются администратором в общую базу, которая доступна для ознакомления всем пользователям и которая используется большинством врачей для диагностики.

Задачи, решаемые с помощью банка знаний пользователями различных типов, и соответствующее им программное наполнение

Как отмечалось в разделе 1, к программному наполнению разрабатываемого банка знаний относятся программные средства трех типов: редакторы информационного наполнения; средства по интеллектуальной обработке ИН (решатели задач); административная подсистема.

Пользователи банка знаний по медицинской диагностике делятся на следующие группы: обслуживающие пользователи (администратор); носители информации (эксперты); прикладные пользователи (лечащие врачи и гости).

Администратор обеспечивает работу пользователей с банком, выполняя следующие функции:

- предоставление полномочий пользователям и создание разделов и баз в ИН;

- интеграция формируемых экспертами знаний о заболеваниях в общую базу;

- выполнение оптимизации общей базы знаний о заболеваниях и сохранение всей информации о заболеваниях, включая данные оптимизации, на кластер – для использования врачами при диагностике.

Для решения первой задачи администратору необходима административная подсистема. Ведение учетной записи пользователя включает ее создание, удаление и задание полномочий: к каким частям ИН и ПН имеет доступ пользователь и каков этот доступ. Некоторые части ИН можно только просматривать, другие – редактировать. Некоторые программы из ПН можно запускать, другие – нет.

Для выполнения второй задачи администратор использует подсистему интеграции знаний, с помощью которой он осуществляет контроль за тем, как пользователи (эксперты) ведут наполнение индивидуальных баз знаний в разделах ИН банка. Он имеет возможность включать элементы

индивидуальных баз знаний в общую базу знаний.

Программное средство, выполняющее оптимизирующее преобразование базы знаний о заболеваниях, работает по алгоритму, описанному в [6]. Оно анализирует информацию из базы знаний о заболеваниях и определяет, при каких заболеваниях может наблюдаться каждое значение каждого признака. Получив данные оптимизации, это же программное средство переносит на кластер всю информацию из базы знаний о заболеваниях, включая результаты оптимизации. Этот процесс назовем публикацией базы знаний, а само средство – *публикатором*.

Эксперты банка занимаются формированием базы наблюдений и индивидуальных баз знаний о заболеваниях в каждом разделе ИН банка. Для этого в ПН банка знаний присутствуют два редактора – для базы наблюдений и для базы знаний о заболеваниях.

Помимо занесения знаний о заболеваниях в базу, эксперт также имеет возможность протестировать ее. Для этого он должен воспользоваться сначала упомянутым выше *публикатором* для оптимизации персональной базы и ее трансляции на кластер, а затем – подсистемой диагностики.

Если в процессе деятельности по формированию базы знаний о заболеваниях эксперт получает существенные результаты, которыми необходимо пополнить общую базу знаний, он посылает запрос на пополнение ИН администратору, который принимает решение о возможности ее добавления в банк посредством административной системы и выполняет это.

Прикладные пользователи делятся на врачей и гостей.

Врачи могут формировать истории болезней пациентов в общей базе и запускать диагностику для любого из пациентов. Для этого в ПН банка присутствует редактор базы историй болезней, с помощью которого врач заносит общую информацию о больном, значения наблюдений из истории жизни, истории заболевания и данные осмотров – в соответствии с моделью истории болезни и с использованием наполнения базы наблюдений.

Кроме того, врачи пользуются присутствующим в программном наполнении средством, выполняющим медицинскую диагностику, алгоритм которой описан в [6]. При работе с этой системой врач выбирает историю болезни, запускает для нее процесс диагностики, после чего получает результаты.

Архитектура этого программного средства несколько сложнее, чем архитектура других компонентов программного наполнения. Поскольку сам алгоритм диагностики реализуется как параллельное приложение и выполняется на кластере, то сверху он снабжается еще двумя компонентами: интерфейсом пользователя (работает на ПК пользователя) и подсистемой, работающей на сервере банка знаний. Эта подсистема выполняет последовательно три действия – пересылку истории болезни на кластер, запуск параллельного алгоритма на кластере для обработки этой истории болезни, трансляцию результатов диагностики с кластера пользователю.

Также врачи могут в каждом разделе банка просматривать базу наблюдений и общую базу знаний о заболеваниях в ознакомительных целях.

Гости имеют возможность только знакомиться с содержанием базы наблюдений и общей базы знаний о заболеваниях, находящихся в любом разделе информационного наполнения банка. Для этого они используют соответствующие редакторы, запускаемые в режиме «только чтение».

Суммируя сказанное, в программном наполнении банка знаний по медицинской диагностике необходимо разработать следующие семь программных средства:

- редактор базы наблюдений;
- редактор базы знаний о заболеваниях;
- редактор базы историй болезней пациентов;
- подсистема интеграции знаний;
- административная подсистема;
- публикатор базы знаний о заболеваниях – средство, выполняющее оптимизацию базы знаний о заболеваниях и перенос всех знаний о заболеваниях на кластер;
- средство, выполняющее медицинскую диагностику.

Заключение

В работе предложена концепция *банка знаний по медицинской диагностике*. Он состоит из информационного и программного наполнений. Базы знаний предметной области и данных действительности в информационном наполнении строятся в соответствии с моделью онтологии предметной области, состоящей из трех частей (модель наблюдений, модель знаний о заболевании, модель истории болезни пациента). В программное наполнение банка входят средства редактирования баз знаний и данных, решатель задачи медицинской диагностики (часть которого работает на кластере), программное средство, выполняющее оптимизацию базы знаний о заболеваниях и ее трансляцию на кластер (*публикатор*), и административная подсистема.

Реализация данного банка возможна в рамках проекта Многоцелевого Банка Знаний, разработанного в ИАПУ ДВО РАН [10,11], предназначенного для поддержки жизненного цикла совместимых систем обработки информации.

ЛИТЕРАТУРА

1. Москаленко Ф.М. Задача медицинской диагностики и алгоритм ее решения, допускающий распараллеливание // Информатика и системы управления. – 2005. – № 2(10). – С.52-63.
2. Detmer W.M., Shortliffe E.H. Using the Internet to Improve Knowledge Diffusion in Medicine. 1997. http://www-smi.stanford.edu/pubs/SMI_Reports/SMI-97-0658.pdf.

3. Зильбер А.П., Шифман Е.М., Павлов А.Г., Белоусов С.Е. Интернет-проект "Компьютерная диагностика преэклампсии", 1998. <http://critical.onego.ru/critical/medlogic/>
4. Клещев А.С., Москаленко Ф.М., Черняховская М.Ю. Модель онтологии предметной области "Медицинская диагностика". Часть 1. Неформальное описание и определение базовых терминов. // НТИ. Сер. 2. – 2005. – №12. – С.1-7.
5. Клещев А.С., Москаленко Ф.М., Черняховская М.Ю. Модель онтологии предметной области "Медицинская диагностика". Ч.2. Формальное описание причинно-следственных связей, причин значений признаков и причин заболеваний // НТИ. Сер. 2. – 2006. – №2. – С.19-30.
6. Москаленко Ф.М. Параллельный оптимизированный алгоритм медицинской диагностики // Информатика и системы управления. – 2006. – №1(11). – С.87-98.
7. Москаленко Ф.М. Экспериментальное исследование временной сложности параллельного алгоритма диагностики, основанного на реальной онтологии медицины // Информатика и системы управления. – 2006. – №2(12). – С.42-53.
8. Прохоров О.В. Клиент-серверная экспертная система для телемедицины. Международная конференция молодых ученых по математическому моделированию и информационным технологиям. 29-31 октября 2002г., Новосибирск, Академгородок. <http://www.ict.nsc.ru/ws/YM2002/4686>
9. Sterling W.M. A Massive Repository for the National Medical Knowledge Bank, Proceedings of the Eighteenth IEEE Symposium on Mass Storage Systems and Technologies – Vol. 00, 2001. – P. 191. <http://storageconference.org/2001/2001CD/15sterli.pdf>.
10. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Многоцелевой банк знаний. // Информационные технологии. – 2006. – №2. – С.2-8.
11. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Требования к многоцелевому банку знаний // Информационные технологии. – 2006. – №4. – С.21-28.

Статья представлена к публикации членом редколлегии А.С. Клещевым.