

## КОНЦЕПЦИЯ КОМПЬЮТЕРНОГО БАНКА ЗНАНИЙ ПО МЕДИЦИНСКОЙ ДИАГНОСТИКЕ

Филипп Москаленко, Александр Клещёв, Мери Черняховская

**Аннотация:** В работе описано информационное и программное наполнение компьютерного банка знаний по медицинской диагностике. Определены классы его пользователей и задачи, ими решаемые. Информационное наполнение специализированного банка знаний содержит три онтологии: онтологию наблюдений в области медицинской диагностики, онтологию базы знаний (заболеваний) по медицинской диагностике и онтологию историй болезни, а также три класса информационных ресурсов в различных областях медицины – базы наблюдений, базы знаний, и базы данных (пациентов), соответствующие этим онтологиям. Программное наполнение содержит редакторы информации (онтологий, баз наблюдений, знаний и данных), а также решатель задачи медицинской диагностики.

**Keywords:** Medical Diagnostics, ontology model, parallel computing, knowledge bank.

**ACM Classification Keywords:** I.2.1 Applications and Expert Systems, J.3 Life and Medical Sciences.

---

### Введение

---

Одним из приложений систем искусственного интеллекта являются системы медицинской диагностики. Их применение помогает врачу повысить качество своей работы. Задачей таких систем является определение заболеваний (одного или нескольких), которыми возможно болен пациент, на основе данных о его наблюдениях. Необходимыми компонентами таких систем являются подсистема доверия, которая показывает пользователям-врачам, какими знаниями обладает система диагностики, и подсистема объяснения, которая разъясняет пользователю, на основе каких рассуждений и знаний системой предлагаются те или иные решения.

В настоящее время для решения задачи медицинской диагностики разрабатываются два класса систем, различающиеся методами, которые положены в их основу. Один класс составляют системы, основанные на статистических и других математических моделях – их основой служат математические алгоритмы, занимающиеся поиском обычно частичного соответствия между симптомами очередного пациента и симптомами наблюдавшихся ранее пациентов, диагнозы которых известны [1 – 4]. Однако такие системы не имеют подсистем доверия и средств формирования понятного врачам объяснения полученного результата.

Системы второго класса основаны на знаниях экспертов. В них алгоритмы оперируют информацией о пациенте и знаниями о заболеваниях, представленными в форме, в той или иной степени приближенной к представлениям врачей (и описанных экспертами-врачами), что достигается за счёт явного или неявного использования онтологий медицинской диагностики. Именно в системах такого типа возможно создание подсистем доверия, а также построение компонента объяснения, способного дать врачу результаты анализа данных пациента, которые привели к полученному решению задачи.

Используемые в таких системах модели онтологии учитывают изменение значений признаков во времени [1], связи симптомов и заболеваний (например, с помощью логических правил) [2, 5], деление наблюдений на несколько групп (например, клинические, лабораторные, морфологические данные в [4]), представление состояния пациента в виде многоуровневой модели [6].

Алгоритмы в таких системах пытаются имитировать ход рассуждений врача при постановке диагноза [7, 8], занимаются поиском соответствия информации о больном и клинических картин заболеваний, описанных врачом [9], либо обработкой заданных экспертами правил, описывающих связи наблюдений и заболеваний [5, 8, 10, 11].

Анализ разрабатываемых в последние годы систем медицинской диагностики второго класса показал, что используемые в них онтологии медицинской диагностики являются сравнительно простыми и одновременно не отражают такие повсеместно используемые врачами в их практике знания предметной области как: знания о причинах заболеваний; знания о различных типах причинных связей между признаками и заболеваниями; знания о воздействиях событий на значения признаков при заболеваниях и у здоровых пациентов; знания о различных вариантах изменения значений признаков, зависящие от анатомо-физиологических особенностей пациентов.

Помимо этого, одним из отрицательных свойств некоторых разрабатываемых систем является 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 определённого медицинского учреждения и не доступны 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 презеклампсии [12], 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 приведённые выше особенности медицинских знаний, в которой их модель имеет форму наиболее близкую к представлениям экспертов и позволяет определять не только диагноз пациента, но и объяснять его. Такая система должна проводить диагностику за приемлемое для врача время (несмотря на то, что в её 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.

Работа выполнена в рамках конкурсного проекта ДВО РАН № 06-III-A-01-457 «Проектирование, разработка и развитие банка медицинских знаний в сети Интернет».

---

## **1. 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 были разработаны несколько моделей онтологии медицинской диагностики. Некоторые из этих моделей применялись при разработке систем диагностики, основанных на знаниях экспертов. Как отмечалось во введении, каждая из них в чём-то была лучше, а в чём-то хуже других [1, 2, 4, 5, 6]. Для объединения преимуществ этих моделей и создания модели онтологии, наиболее приближенной к представлениям знаний в области медицинской диагностики была разработана и описана онтология и её модель [13,14,15]. 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.

Соотношения между неизвестными и параметрами делятся на следующие смысловые группы:

- 1) Соотношения между знаниями о причинно-следственных отношениях и причинно-следственными связями, имеющими место в ситуации.
- 2) Соотношения, определяющие причинно-следственные связи, являющиеся причинами значений каждого признака на разных интервалах времени.
- 3) Соотношения, определяющие для каждого признака свойства границ интервалов разбиения оси времени, связанного с этим признаком.
- 4) Соотношения, определяющие причину для каждого заболевания, входящего в диагноз.

На основе данной модели онтологии в работе [16] поставлена общая задача медицинской диагностики: определить возможные диагнозы больного на основе знаний предметной области и данных его обследования, к которым относятся значения признаков (в моменты их наблюдения), значения его анатомо-физиологических особенностей (постоянные во времени) и значения произошедших с ним событий (в моменты, когда они происходили), а также для каждого такого диагноза сформировать его причину (некоторое событие или другие заболевание) и объяснение (путём указания причин наблюдаемых значений признаков).

В связи с тем, что приведённая модель онтологии учитывает большое число связей между процессами, происходящими в организме больного, можно ожидать, что алгоритм решения сформулированной выше задачи медицинской диагностики, анализирующий все эти связи, будет иметь высокую вычислительную сложность. Одним из путей повышения эффективности такого алгоритма является его распараллеливание и выполнение на многопроцессорной вычислительной системе. Наибольшего эффекта от распараллеливания можно достичь при решении некоторой частной, но тем не менее практически важной задачи, при решении которой признаки можно анализировать независимо друг от друга и, таким образом, параллельно на узлах многопроцессорного вычислительного комплекса. Такая задача получается при наложении на используемую модель онтологии нескольких ограничений: пациент может быть либо здоров, либо болен только одним заболеванием, а каждое заболевание имеет ровно один период развития.

В работе [16] приводится постановка задачи в терминах упрощённой онтологии и алгоритм её решения, а в работе [17] 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 [17] приведены в работе [18]. Они показали, что:

- а) использование оптимизирующего преобразования базы знаний о заболеваниях заметно ускоряет процесс постановки диагноза, в особенности – при наличии большого количества заболеваний в базе знаний;
- б) наибольшая скорость постановки диагноза наблюдается при выполнении не менее одного и не более двух процессов на каждом узле кластера;
- в) время работы алгоритма на всех использованных тестовых наборах данных при запуске оптимального числа процессов и использовании оптимизации базы знаний не превышает нескольких минут.

Описанная онтология и основанный на ней алгоритм медицинской диагностики могут быть использованы при создании системы, которая должна обеспечивать процесс согласованного решения комплекса задач по сбору, формализации, переводу в машиночитаемое представление, инженерии, хранению, управлению и обработке данных и знаний в области медицинской диагностики и является объединением всей этой информации в единый ресурс с возможностью удаленного доступа к нему многим пользователям.

Разрабатываемый ресурс назовем Банком знаний по медицинской диагностике. Для реализации указанных требований, архитектура этой системы должна соответствовать парадигме трёхзвенного программного обеспечения, т.е. банк знаний должен состоять из следующих трёх частей:

- информационного наполнения с некоторым стандартным интерфейсом доступа к хранимой информации и унифицированным форматом хранения этой информации;
- программного наполнения, ориентированного на интеллектуальную поддержку пользователей банка и включающего: средства по редактированию данных из информационного наполнения, средства по их обработке (оптимизация базы знаний о заболеваниях и медицинская диагностика пациента) и административную подсистему;
- интерфейсов к программным средствам из программного наполнения.

---

## **2. Информационное наполнение Банка знаний по медицинской диагностике**

---

В информационном наполнении банка в первую очередь должна храниться модель онтологии предметной области медицинская диагностика [13,14,15]. В соответствии с ней в банк заносятся знания предметной области (описание заболеваний) и данные пациентов.

Модель онтологии, хранящаяся в информационном наполнении специализированного банка знаний, должна состоять из трёх частей:

- модель наблюдений,
- модель знаний о том, как заболевания влияют на значения признаков (модель знаний о заболеваниях),
- модель пациента (модель истории болезни пациента).

Модель онтологии базы наблюдений описывает структуру наблюдений и их значений. К наблюдениям относят наблюдаемые признаки, события и анатомо-физиологической особенности. На основе этой модели строится база наблюдений, куда входит описание названий событий, признаков и особенностей, а также перечисление всех их возможных значений.

Модель онтологии базы пациентов строится на основе модели онтологии базы наблюдений и задаёт структуру описания состояние пациента во времени. События могут происходить с пациентом в различные моменты времени и иметь различные значения. Признаки также в различные моменты времени могут иметь различные значения, а анатомо-физиологические особенности пациента имеют постоянные во времени значения. На основе данной модели пользователи, выполняющие медицинскую диагностику, формируют базу историй болезней пациентов, т.е. создают в информационном наполнении записи о пациентах: наблюдаемых у них в различные моменты времени значениях признаков, происходящих с ними событий и значения анатомо-физиологических особенностей.

В модели знаний о заболеваниях должны быть описаны в терминах модели наблюдений основные термины знаний (включая отношения между ними), в том числе заболевания (включая их причины) и нормальное состояние пациента. На основе схемы, задаваемой этой моделью, экспертами описываются

конкретные заболевания (значения признаков при заболеваниях задаются клиническими проявлениями и клиническими проявлениями, изменёнными воздействием событий), их причины (задаются этиологиями) и нормальное состояние пациента (описываются нормальные реакции и реакции на воздействие событий).

На рис. 1 представлена общая схема информационного наполнения банка знаний по медицинской диагностике. Стрелки указывают, как один компонент информационного наполнения используется при формировании другого.

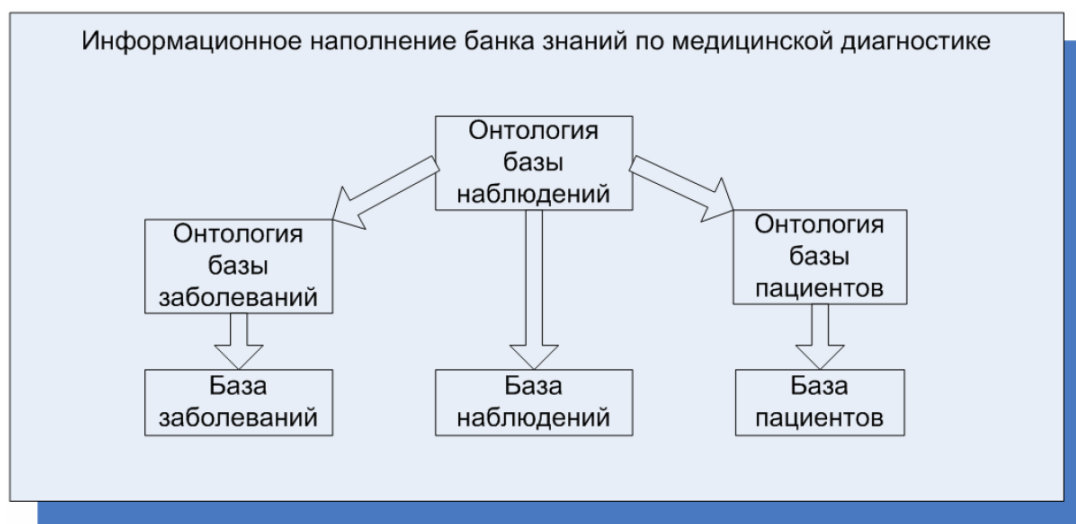


Рис. 1. Информационное наполнение банка знаний по медицинской диагностике.

### 3. 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 трёх типов:

- редакторы информационного наполнения,
- решатели задач,
- административная система.

В соответствии с приведёнными в разделе 2 составляющими информационного наполнения банка знаний по медицинской диагностике необходимо разработать следующие редакторы информационного наполнения:

- редактор онтологии наблюдений;
- редактор онтологии базы заболеваний (описания значений признаков при заболеваниях и у здорового пациента);
- редактор онтологии историй болезней пациентов;
- редактор базы наблюдений;
- редактор базы знаний о заболеваниях;

- редактор базы историй болезней пациентов.

При помощи редакторов онтологий на начальном этапе реализации банка знаний задаются модели соответствующих онтологий. Впоследствии данные средства не предполагается использовать, так как задание баз знаний необходимо вести в рамках жёстко закреплённых моделей онтологий.

В качестве алгоритма для решателя задачи медицинской диагностики используется упомянутый выше алгоритм [17], учитывающий все связи между знаниями и действительностью, заложенные в модели онтологии [14,15].

Программное средство, выполняющее оптимизирующее преобразование базы знаний о заболеваниях, должно работать по алгоритму, описанному в [17]. Оно анализирует информацию из базы заболеваний и определяет, при каких заболеваниях может наблюдаться каждое значение каждого признака. На основе этой информации алгоритм диагностики при формировании гипотез о заболевании формирует некоторое множество гипотез, возможно, меньшее, чем количество всех заболеваний в базе.

Всех пользователей банка знаний в области медицинской диагностики делятся на:

- обслуживающих пользователей (администратор и инженер-программист),
- носителей информации (экспертов-врачей),
- прикладных пользователей (врачей-диагностов и гостей).

Администратор следит за функционированием всего банка знаний при помощи административной подсистемы, выполняющей две функции: работа с пользовательскими записями и контроль над деятельностью пользователей.

Инженер-программист ответственен за функциональность всех компонент банка знаний. Он выполняет подготовку банка к функционированию и устраняет сбои в его работе, при возникновении таковых.

Врачи-эксперты занимаются редактированием баз наблюдений и знаний о заболеваниях, выполняя это в соответствии с моделью соответствующих онтологий. Такой пользователь для индивидуальной работы может создавать свою копию ИН или его части, где он может производить любые изменения, необходимые ему для работы. Если в процессе своей исследовательской или экспериментальной деятельности он получил существенные результаты, пополняющие базу данных, то он посылает запрос на пополнение информационного наполнения администратору, который принимает решение о возможности ее добавления в банк посредством административной системы.

Прикладные пользователи делятся на врачей, работающих с базой историй болезней и проводящих диагностику, и гостей, способных только знакомиться с содержанием базы наблюдений и базы знаний о заболеваниях, находящихся в информационном наполнении банка.

---

#### **4. Реализация Банка знаний по медицинской диагностике**

---

Реализация требуемой системы может быть выполнена с использованием разработанного в ИАПУ ДВО РАН Многоцелевого Банка Знаний (МБкЗ) [19,20], предназначенного для поддержки жизненного цикла совместимых систем обработки информации.

В таком случае, программное наполнение Банка знаний по медицинской диагностике включает в себя:

- редактор ИРУО, являющийся частью МБкЗ, который используется для редактирования на языке ИРУО моделей онтологий и целевой информации, формируемой по этим моделям [21,22];

- административную систему, также входящую в состав МБкЗ, с помощью которой можно выполнять управление пользователями;
- решатель задачи медицинской диагностики;
- преобразователь базы знаний о заболеваниях.

Алгоритм диагностики реализован в виде параллельного приложения на языке С++, которое выполняется на многопроцессорном вычислительном комплексе (кластере) под управлением ОС Linux. Доступ к данному ресурсу предоставляется серверу банка знаний. На сервере банка знаний работает слабо-функциональная серверная часть решателя, реализованная на языке Java, которая ведёт общение с пользователем через интерфейсную часть и выполняет запуск диагностики на МВС с последующей трансляцией её результатов к клиенту-пользователю.

Поскольку алгоритм диагностики для быстрой постановки диагноза должен иметь быстрый доступ к знаниям и данным с как можно меньшим количеством промежуточных трансляционных узлов/прослоек, то до начала диагностики необходимо отправить на кластер наполнение базы знаний о заболеваниях, для чего в подсистему оптимизации этой базы знаний встроены выполняющий это код. При начале диагностики, из информационного наполнения банка на кластер транслируется лишь история болезни пациента.

По окончании диагностики, все полученные результаты (опровергнутые диагнозы, подтверждённые диагнозы с их причинами и объяснениями значений наблюдаемых признаков) транслируются с кластера сервер-частью решателя в интерфейсную клиент-часть и отображаются на экране пользователя.

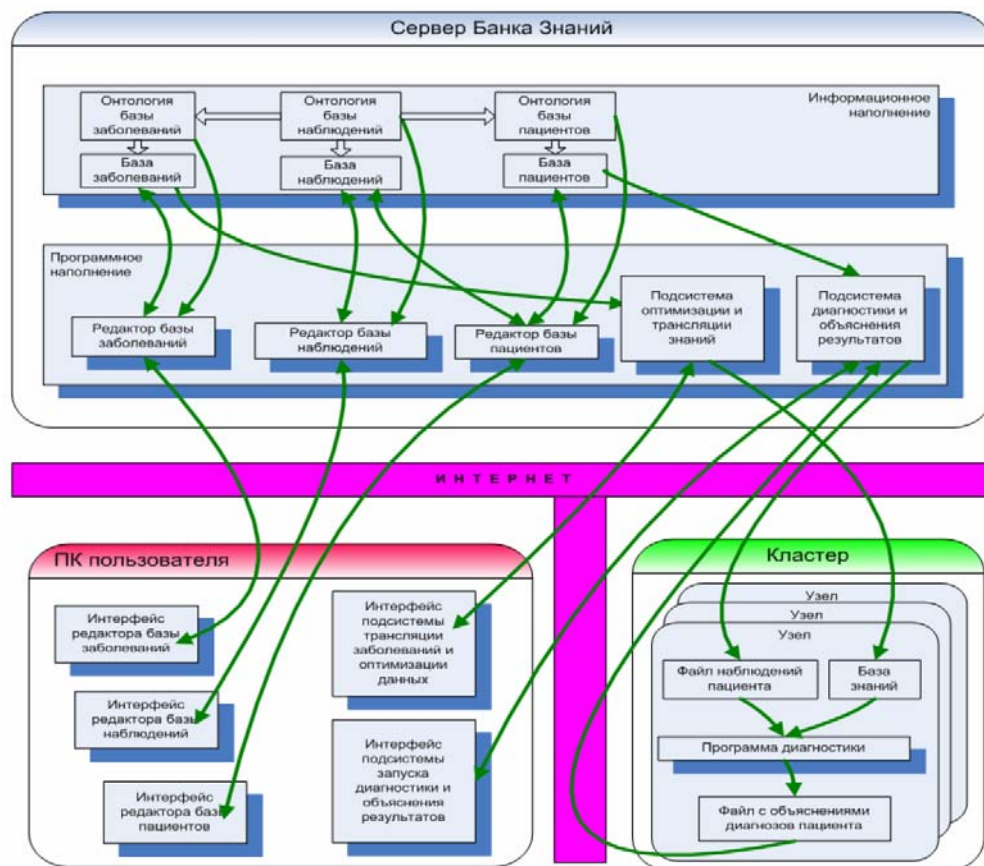


Рис.2. Архитектура банка знаний по медицинской диагностике.

На рис. 2 представлена общая архитектура банка знаний по медицинской диагностике, выполненная в рамках концепции СБкЗ. Ввиду сложности схемы на ней не отображается административная оболочка и её интерфейс, включая хранилище информации о пользователях банка, а также редакторы онтологий (в данном случае – Редактор ИРУО), так как они используются только на начальном этапе существования банка для задания моделей этих онтологий. Кроме того, следует понимать, что в качестве редактора баз знаний также используется Редактор ИРУО, но для каждой базы может быть разработан и включён в состав программного наполнения банка самостоятельный редактор, более удобный пользователям.

---

### Заключение

---

В работе предложена концепция банка знаний по медицинской диагностике. Он содержит модель онтологии предметной области, состоящую из трёх частей (модель наблюдений, модель знаний о заболевании, модель истории болезни пациента), знания предметной области (база наблюдений и база знаний о заболеваниях) и данные действительности (база историй болезней пациентов). Кроме того, в банк входит средство редактирования моделей онтологий, баз знаний и данных, решатель задачи медицинской диагностики, программное средство, выполняющее оптимизацию базы знаний о заболеваниях, и административная подсистема. Предложены пути реализации данного банка в рамках проекта Многоцелевого Банка Знаний, разработанного в ИАПУ ДВО РАН.

---

### Список литературы

---

1. Генкин А.А. О последовательной стратегии Байеса и механизме принятия решений в интеллектуальной системе ОМИС. 1998. [http://intels.spb.ru/english\\_med\\_informatiks-2.html](http://intels.spb.ru/english_med_informatiks-2.html)
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](http://www-smi.stanford.edu/pubs/SMI_Reports/SMI-97-0658.pdf)
3. Алексеев А.В. Применение методов информатики и компьютерной техники в дифференциальной диагностике аппендицита и аппендикулярной колики. <http://arkadyal.chat.ru/programs.htm>
4. Бурдаев В.П. Прикладные экспертные системы, основанные на оболочке "TECHEXP". 2002. <http://uacm.kharkov.ua/eng/index.shtml?eexpert.htm>
5. ООО "Админ". Экспертная система "Диагностика коматозных состояний". 1992. [http://www.adminru.com/start\\_e.htm](http://www.adminru.com/start_e.htm)
6. Chabat F., Hansell D.M., Guang-Zhong Yang. Computerized Decision Support in Medical Imaging. 1997. <http://www.doc.ic.ac.uk/~gzy/pub/chabat-decision-support.pdf>
7. Гаськов А.П., Валивач М.Н. Экспертная система медицинской диагностики "Консилиум". 2000. [http://www.eksi.kz/consilium/librar/esmd\\_short.htm](http://www.eksi.kz/consilium/librar/esmd_short.htm)
8. Lhotska L., Vlcek T. Efficiency enhancement of rule-based expert systems. Proceedings of the 15th IEEE symposium on computer-based medical systems (CBMS 2002). pp.53-58.
9. Matsumoto T., Ueda Y., Kawaji Sh. A software system for giving clues of medical diagnosis to clinician. Proceedings of the 15th IEEE symposium on computer-based medical systems (CBMS 2002). pp. 65-58.
10. Filho M.M., Palombo C.R., Sabbatini R.M. TMJ Plus: A Knowledge Base and Expert System for Diagnosis and Therapeutics of the Temporomandibular Joint Disorders. 1994. <http://www.epub.org.br/ojdom/vol03n03.htm>
11. Sawar S.J. Diagnostic Decision Support System of POEMS. 1999. <http://www.cbl.leeds.ac.uk/sawar/projects/poems/poems-in-detail.html>
12. Зильбер А.П., Шифман Е.М., Павлов А.Г., Белоусов С.Е. Интернет Проект "компьютерная диагностика преэклампсии". 1998. <http://critical.onego.ru/critical/medlogic/>



13. Chernyakhovskaya M.Yu., Kleschev A.S., Moskalenko F.M. A metaontology for medical diagnostics of acute diseases. Part 1. An informal description and definitions of basic terms. IJ ITA, 2007. (в печати)
14. Chernyakhovskaya M.Yu., Kleschev A.S., Moskalenko F.M. A metaontology for medical diagnostics of acute diseases. Part 2. A formal description of cause-and-effect relations. IJ ITA, 2007. (в печати)
15. Chernyakhovskaya M.Yu., Kleschev A.S., Moskalenko F.M. A metaontology for medical diagnostics of acute diseases. Part 3. A formal description of the causes of signs' values and of diseases. IJ ITA, 2007. (в печати)
16. Москаленко Ф.М. Задача медицинской диагностики и алгоритм ее решения, допускающий распараллеливание // Информатика и системы управления. – 2005. – № 2(10). – С.52-63.
17. Москаленко Ф.М. Параллельный оптимизированный алгоритм медицинской диагностики // Информатика и системы управления. – 2006. – № 1(11). – С.87-98.
18. 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.
19. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Многоцелевой банк знаний. // Информационные технологии – 2006. – №2. – С.2-8.
20. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Требования к Многоцелевому банку знаний. // Информационные технологии – 2006. – №4. – С.21-28.
21. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Универсальный подход к решению проблемы редактирования информации. // Информационные технологии – 2006. – №5. – С.25-31.
22. Орлов В.А., Клещев А.С. Компьютерные банки знаний. Модель процесса редактирования информационного наполнения. // Информационные технологии – 2006. – №7. – С.11-16.

---

### Информация об авторах

---

Москаленко Ф.М. - [philipmm@yahoo.com](mailto:philipmm@yahoo.com)

Клещев А.С. - [kleschev@iacp.dvo.ru](mailto:kleschev@iacp.dvo.ru)

Черняховская М.Ю. - [chernyah@iacp.dvo.ru](mailto:chernyah@iacp.dvo.ru)

Институт автоматизации и процессов управления, Дальневосточное отделение Российской Академии Наук, Российская Федерация, г. Владивосток, ул. Радио, д.5.