

Технология МКЛАСС для построения полных и непротиворечивых баз знаний в задачах медицинской диагностики

Болотов А.А., Барчуков В.Г.
ООО Радифенс, Московская область

Создание компьютерных систем, имитирующих искусство эксперта или компьютерной копии высококвалифицированного врача до сих пор является актуальной задачей. При построении таких систем одной из центральных является задача передачи знаний эксперта компьютерной программе, известная как проблема приобретения знаний.

Вопрос о выявлении и передаче компьютеру человеческих знаний и умений является крайне сложным. На пути решения этой задачи стоит множество препятствий:

1. Человек не может сообщить абстрактных правил, которыми он руководствуется, решая ту или иную конкретную задачу. Человеческие умения чаще всего хранятся на подсознательном уровне.

2. В любой области человеческой деятельности имеется большое количество (десятки и сотни тысяч) возможных практических ситуаций для принятия решений, при анализе которых проявляются умения человека. Желательно уметь передавать эти возможности ЭВМ.

3. Люди, передающие компьютеру в том или ином виде свои знания, неизбежно ошибаются.

4. Процессы извлечения знаний отличаются высокой трудоемкостью и большими интеллектуальными нагрузками, приходящимися на эксперта и инженера по знаниям, связанными с вербализацией экспертизы, управлением процессом коммуникации и необходимостью усвоения, анализа и документирования больших объемов новой информации.

Важнейшей характеристикой процедуры извлечения знаний является ее психологическая корректность. Поскольку основным источником информации является человек, необходимо принимать во внимание особенности переработки им информации. При этом следует отметить, что ее достоверность существенно зависит от двух факторов: понятности и привычности операций, выполняемых экспертом при извлечении знаний, и соответствии операций по переработке информации возможностям человека.

Известно, что количество психологически корректных операций получения информации в многоаспектных задачах сравнительно невелико. Но только эти способы ведут к надежным базам знаний (БЗ) большого объема. Кроме этого, существуют ограничения по параметрам решаемой задачи. Люди ведут себя надежно и непротиворечиво при определенных значениях количества классов решений, числа признаков и числа возможных значений для каждого из них. Эти ограничения важно знать и использовать при разработке методов извлечения экспертных знаний.

Еще одной важной характеристикой, связанной с получением информации от экспертов, является полнота извлечения и представления

знаний в компьютерной системе, а также непротиворечивость (безошибочность) базы знаний. Под полнотой и непротиворечивостью базы знаний, в общем смысле, понимают способность системы формировать непротиворечивое решение для любой ситуации в рамках ограниченной предметной области.

Среди источников неполноты и ошибочности базы знаний могут быть выделены следующие два фактора:

1. Знания и умения эксперта на когнитивном уровне, отражающие его квалификацию и степень проработанности проблемной области, могут быть неполными и содержать ошибочные представления о реальности;

2. При передаче знаний компьютеру могут появляться ошибки, связанные с неточностью и неоднозначностью естественного профессионального языка эксперта, так и с вербализацией тех или иных представлений эксперта.

С учетом вышеизложенного нами была разработана операционная технология построения больших, полных и непротиворечивых баз знаний (содержащих десятки тысяч решающих правил экспертов), основанная на максимальном использовании всех "островков знаний" рассматриваемой предметной области.

Основные характеристики этой технологии:

1. Возможность использования базы фактографических данных (подобранные истории болезни или специальные карточки, базы данных и т.п.);

2. Для создания базы знаний используется психологически корректная процедура опроса эксперта, при которой эксперт решает привычную для себя процедуру диагностики случая, предъявляемого компьютером по определенному алгоритму;

3. Использование расширенной информации о признаках и структуре задачи (сведения о несовместных признаках, оценки характерности признаков, эвристические правила эксперта и т.п.);

4. Для ускорения создания БЗ система опроса эксперта пытается смоделировать ситуацию таким образом, чтобы предъявлять объекты, находящиеся вблизи границ классов решений.

5. Структура БЗ строится на основе продукций, что позволяет использовать таблицы СУБД для ее хранения и многоаспектного анализа с использованием языка SQL для реляционных баз данных;

6. Возможность анализировать границы классов, выявлять обобщающие решающие правила эксперта, строить деревья решений, оценивать сложность реального случая классификации с точки зрения его близости к граничным объектам и др.

Для поддержки этой технологии была разработана система МКЛАСС с помощью которой было построено более десяти БЗ для задач медицинской диагностики.