

## ПРОГРАММНЫЙ КОМПЛЕКС ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ ХРОНИЧЕСКИХ ВОСПАЛИТЕЛЬНЫХ ЗАБОЛЕВАНИЙ КИШЕЧНИКА

Н.Г.Егорова, В.В. Старовойтов, А.М. Недзьведь  
Объединенный институт проблем информатики, Минск, Беларусь;

*Описана общая схема программного комплекса дифференциальной диагностики хронических воспалительных заболеваний кишечника, основой которого является механизм нечеткого вывода. Подробно рассматривается метод определения весовых коэффициентов нечетких производственных правил.*

### Введение

В настоящее время одной из актуальных проблем гастроэнтерологии является дифференциальная диагностика заболеваний кишечника, в частности, группы хронических воспалительных заболеваний кишечника (ХВЗК), включающей рак толстой кишки, язвенный колит, функциональные расстройства, болезнь Крона, инфекционные заболевания, а также энтеропатию. Каждое из этих заболеваний имеет свои патогенетические особенности и требует определенного подхода к лечению, однако клиническая симптоматика различных заболеваний кишечника во многом схожа, что создает немалые трудности при дифференциальной диагностике. Таким образом, представляется актуальным создание системы, позволяющей обобщить результаты обследований, полученные на различных этапах работы с пациентом, с целью максимального использования этих результатов в процессе дифференциальной диагностики [1].

В силу нечеткости клинической картины ХВЗК, а также субъективизма врача, производящего заключение, при разработке системы диагностики представляется необходимым использование аппарата нечеткой математики, в частности, механизма нечеткого вывода [2]. Вместе с тем, использование указанного механизма требует задания весовых коэффициентов нечетких правил. В предпринятом рассмотрении описана общая схема программного комплекса для дифференциальной диагностики ХВЗК, и особое внимание уделяется подсистеме настройки весовых коэффициентов нечетких производственных правил, в основе которой лежит метрический подход.

### 1. Состав программного комплекса

Постановка диагноза представляет собой процедуру идентификации состояния человека, то есть сопоставление его состояния с другими состояниями, соответствующими как нормальным состояниям, так и нозологическим формам в их различных стадиях развития. Указанное обстоятельство с необходимостью требует предварительного отнесения заболевания человека к некоторой группе заболеваний, в рассматриваемом случае – заболеваний кишечника, с последующей диагностикой одного или нескольких заболеваний рассматриваемой группы, так что программный комплекс диагностики заболеваний кишечника, общая схема которого детально рассматривается в [1], состоит из следующих модулей:

- ввода и редактирования данных о больных;
- оценки и диагностики ХВЗК по каждому из трех наборов данных;
- оценки тяжести ХВЗК по каждому из трех наборов данных;
- суммарной оценки признаков ХВЗК;
- обучения врачей диагностике ХВЗК;
- протоколирования введенных данных.

Работа комплекса ведется с тремя наборами данных:

- клинико-лабораторных исследований;

- эндоскопического обследования;
- морфологических исследований.

Головной модуль ввода и редактирования данных о больных состоит из четырех основных частей: меню, включающего разделы управления данными и помощи по работе с программой; поля основной информации о пациенте, позволяющего просматривать и редактировать введенную информацию; поля с комментариями врача, в котором указываются важные особенности состояния пациента, а также поля управления, в котором определены кнопки запуска модулей программы. Кроме того, программный комплекс включает в себя диагностический блок, в [1] именуемый блоком предварительной диагностики, который производит вычисление коэффициентов близости состояния пациента к состояниям, соответствующим каждому заболеванию группы ХВЗК, а также модуль интерактивного ввода данных, который организует работу с признаками заболевания, и предназначен для отображения дерева признаков, изображенного на рис. 1, которое копируется для каждого пациента. При просмотре признаков пользователем, ветви дерева отмечаются определенным цветом в зависимости от действия пользователя. Пользователь может изменить значение признака или просто его просмотреть.

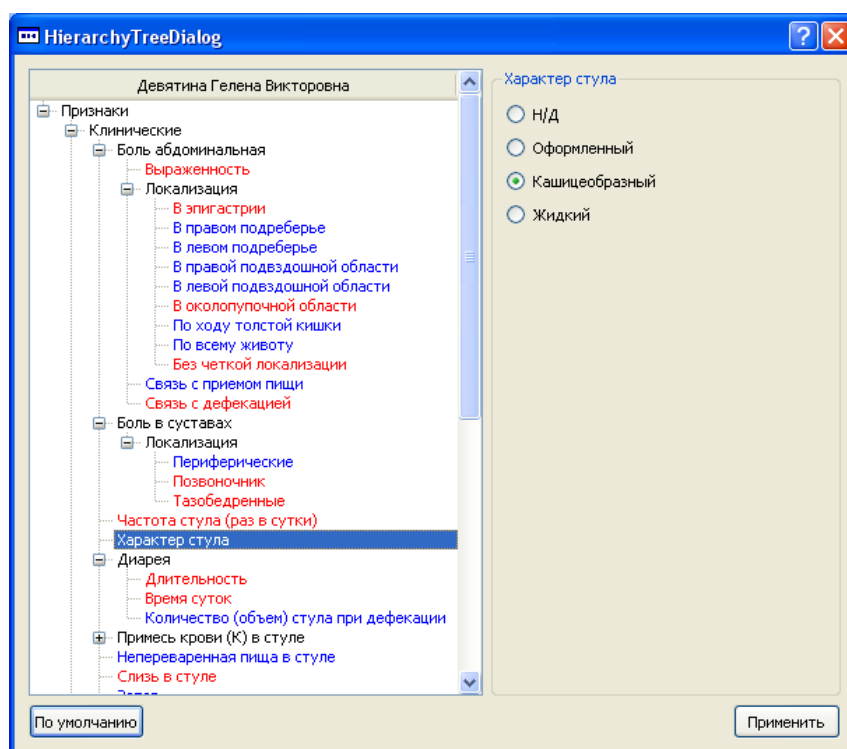


Рис. 1. Копия экрана диалога модуля интерактивного изменения признаков

Для изменения значения признака в правой части экрана реализовано поле, в котором задаются возможные варианты состояния. Реализовано два варианта заполнения:

- троичный, подразумевающий наличие признака, его отсутствие, а также отсутствие соответствующей информации;
- множественный, включающий ряд различных вариантов ответа.

Модуль вычисления, в основе которого основе лежит парсинг дерева признаков и составление формализованных векторов данных, на основе которых блок принимает решения о состоянии больного, организует работу диагностического блока. В результате, в окне предварительной диагностики выводятся коэффициенты близости по каждому типу заболевания.

## 2. Схема взаимодействия диагностического блока с механизмом нечеткого вывода

Основой механизма нечеткого вывода традиционно являются алгоритм Мамдани, алгоритм Ларсена, алгоритм Цукамото или алгоритм Такаги – Сугэно [2]. Этапами нечеткого вывода в каждой из указанных процедур являются:

- формирование базы правил систем нечеткого вывода;
- фаззификация входных переменных;
- агрегирование подусловий в нечетких правилах продукций;
- активизация или композиция подзаключений в нечетких правилах продукций;
- аккумуляирование заключений в нечетких правилах продукций.

Системы нечеткого вывода предназначены для преобразования значений входных переменных в выходные переменные на основе использования нечетких правил продукций. Для этого системы нечеткого вывода содержат базу правил нечетких продукций и реализуют нечеткий вывод заключений на основе посылок или условий, представленных в форме нечетких лингвистических высказываний.

База правил системы нечеткого вывода предназначена для формального представления эмпирических знаний или знаний экспертов в той или иной проблемной области. В системах нечеткого вывода используются правила нечетких продукций, в которых условия и заключения сформулированы в терминах нечетких лингвистических высказываний. В зависимости от количества простых нечетких высказываний в посылке и заключении каждого правила, являющегося элементом базами правил нечетких продукций, выделяют различные типы структуры базы правил. Наиболее общей является так называемая ММО-структура, где правила построены по принципу «много входов – много выходов»:

$$\text{ЕСЛИ } (x^1 = a_1) \text{ И } \dots \text{ (} x^t = a_t \text{) И } \dots \text{ И } (x^m = a_m), \text{ ТО } (y_1 = b_1) \text{ И } \dots \text{ И } (y_g = b_g) (F). \quad (1)$$

В выражении (1) в простом высказывании типа  $x^t = a_t$ ,  $a_t$  – нечеткое множество, являющееся значением лингвистической переменной  $x^t$ , и, не нарушая общности, можно полагать, что  $a_t$  – нечеткий интервал или нечеткое число, и где  $t \in \{1, \dots, m\}$ ;  $F_i, i = 1, \dots, s$  – весовые коэффициенты правил, принимающие значения в интервале от 0 до 1, а  $s$  – число продукций в базе правил. В случае использования в механизме нечеткого вывода правил вида (1), весовые коэффициенты могут быть заданы не только для отдельных правил, но и для каждого из заключений  $y_j, j = 1, \dots, g$  в отдельности.

Определение коэффициентов  $F_i, i = 1, \dots, s$  зачастую производится с помощью экспертных оценок, что является не всегда приемлемым в силу субъективизма экспертов и возможности получения противоречивых индивидуальных оценок, когда коэффициенты определяются путем построения групповой экспертной оценки.

Для вычисления весовых коэффициентов правил можно предложить метрический подход, сущность которого заключается в вычислении расстояния от вектора признаков классифицируемого объекта, до эталонного объекта, с последующим преобразованием значений отдаленности объектов от эталонов в значения близости, для чего, при представлении классифицируемых объектов как нечетких множеств, определенных на универсуме признаков, может быть использована стандартная операция дополнения. Следует указать, что в ряде задач, к примеру, в медицинской диагностике, а также при разработке систем специального назначения эталонные объекты представляются вектором признаков, принимающими значение в интервале, так что каждый признак представляется в виде  $x^t = [x^{t \min}, x^{t \max}]$ ,  $t = 1, \dots, m$ , что позволяет рассматривать каждый эталонный объект в виде нечеткого множества типа 2. Кроме того, признаки объектов могут иметь сложную иерархическую структуру.

Представляя каждый эталонный объект – в рассматриваемом случае это образ того или иного заболевания – как нечеткое множество более высокого, чем первый, типа, представляется возможным обобщить предложенную в [3] процедуру идентификации, основанную на вычислении функции расстояния между эталонным объектом, представленным как нечеткое множество типа 2, и классифицируемым объектом, рассматриваемым как нечеткое множество типа 1, на случай нечетких множеств  $P$  и  $p - 1$  соответственно. К примеру, обобщение предложенной в [3] функции расстояния Хемминга для нечетких множеств типа 2, в случае  $p = 4$ , соответствующего, к примеру, такому признаку, как «боль абдоминальная», как это изображено на рис. 1, будет записываться в виде

$$l_{G_4}(x_i, x_j) = \frac{1}{m_1} \sum_{i_1=1}^{m_1} \left( \frac{1}{m_2} \sum_{u_1, v_1=1}^{m_2} \left( \frac{1}{m_3} \sum_{u_2, v_2=1}^{m_3} \left( \frac{1}{m_4} \sum_{u_3, v_3=1}^{m_4} |a_i^{i_1, u_1, u_2, u_3} - a_j^{i_1, v_1, v_2, v_3}| \right) \right) \right), i, j = 1, \dots, n, \quad (2)$$

где  $m_\ell$  – число признаков  $\ell$ -го уровня градации признака, а  $a_i^{i_1, \dots, i_\ell}$  – соответствующие значения градаций признаков  $i$ -го объекта. Используя предложенный в [1] подход к проектированию систем поддержки принятия решений, заключающийся в одновременном использовании механизма нечеткого вывода и метрического подхода к классификации, представляется целесообразным и методологически оправданным использовать результаты метрической процедуры классификации в качестве весовых коэффициентов правил, как это изображено на рис. 2.

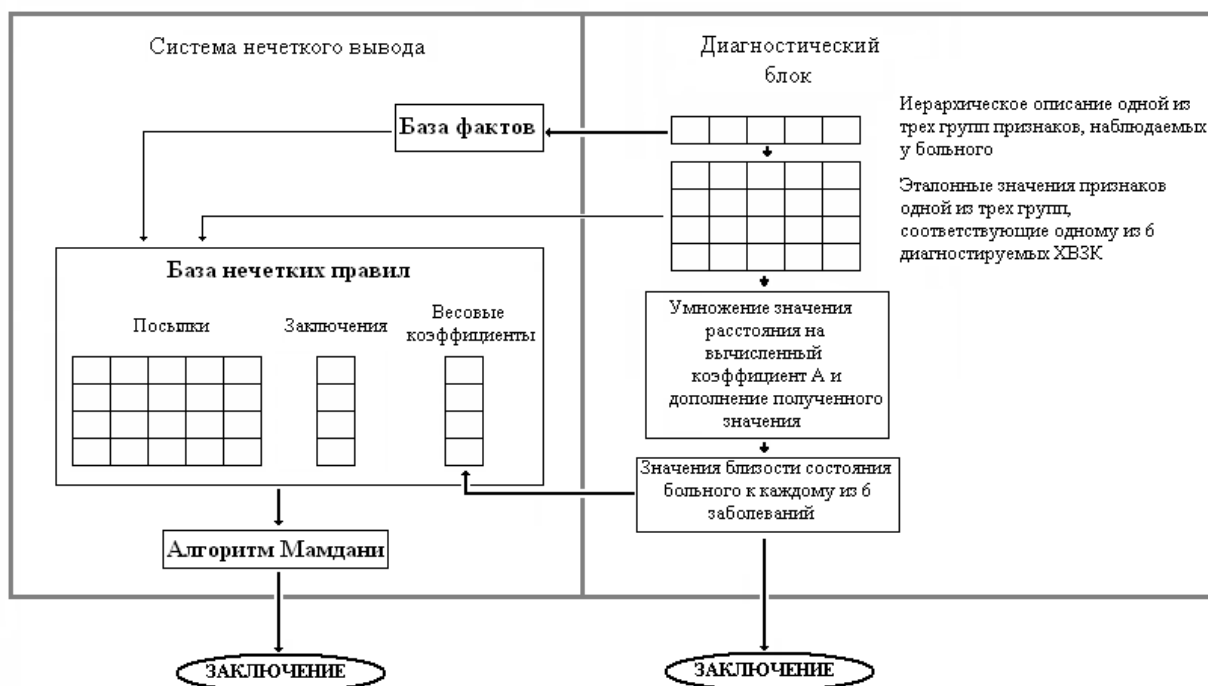


Рис. 2. Настройка весовых коэффициентов и значений лингвистических термов в посылках с помощью диагностического блока

Процедура настройки весовых коэффициентов состоит из 2 шагов:

1. К коэффициентам различия, полученным в результате работы диагностического блока, применяется нормировка, заключающаяся в умножении каждого коэффициента  $d_{*j}$ ,  $j = 1, \dots, 6$  на некоторое зависящее от  $\max_j d_{*j}$  значение  $A$  таким образом, чтобы

выполнялось  $0 < A d_{*j} < 1, \forall j = 1, \dots, 6$ ;

- К каждому нормированному коэффициенту применяется операция дополнения:  $\mu_{*j} = 1 - A d_{*j}, j = 1, \dots, 6$ . Полученные значения  $\mu_{*j}$  будут весовыми коэффициентами правил, в которых устанавливается диагноз  $y_j, j = 1, \dots, 6$ .

Так как работа комплекса ведется с тремя наборами данных, и заключение производится вначале по каждой группе признаков в отдельности, то при выводе окончательного заключения по всем группам признаков одновременно, как это представлено на рис. 3, также используется механизм нечеткого вывода, в котором задействованы нечеткие продукционные правила, входными переменными в которых являются заключения по отдельным группам признаков.

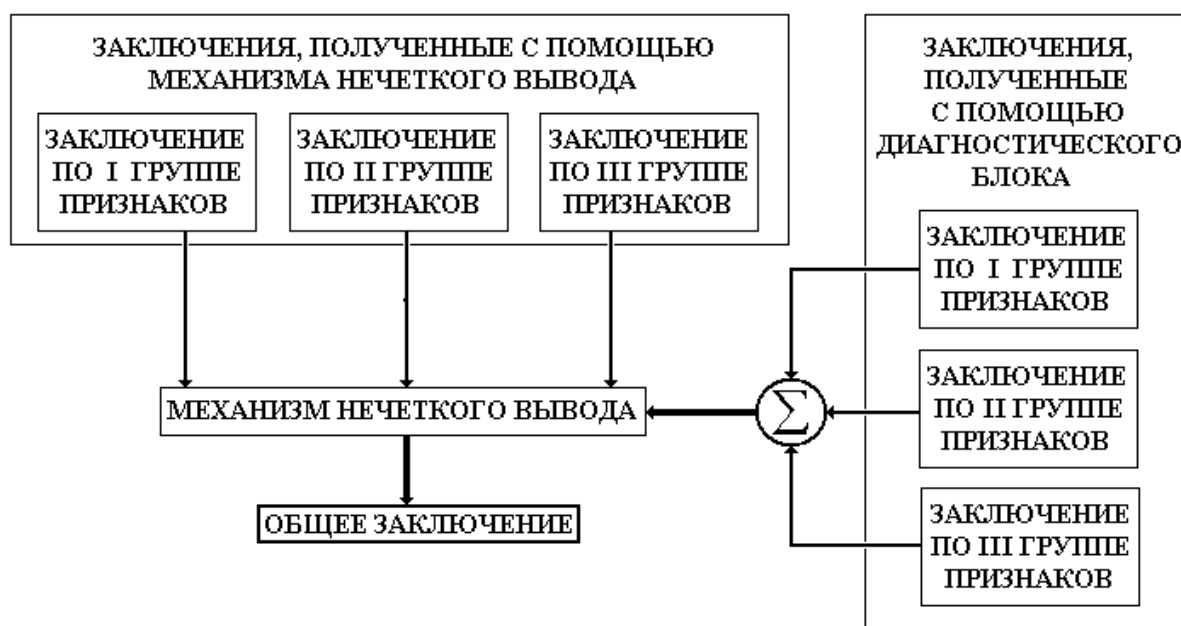


Рис. 3. Схема вывода общего заключения по трем группам признаков

Заключения, полученные с помощью диагностического блока, аккумулируются в сумматоре, изображенном на рис. 3 в виде окружности, обозначенной  $\Sigma$ . Полученные в сумматоре веса заключений по каждому из диагнозов используются в качестве весовых коэффициентов нечетких продукционных правил, по которым производится вывод общего заключения.

## Заключение

Программный комплекс диагностики ХВЗК обладает большой гибкостью в том смысле, что может применяться для диагностики и других групп заболеваний. Разделение функций диагностики между диагностическим блоком и механизмом нечеткого вывода позволяет значительно уменьшить число продукционных правил в механизме нечеткого вывода. Кроме того, реализованный в подсистеме настройки весовых коэффициентов нечетких продукционных правил метрический подход позволяет настраивать весовые коэффициенты каждый раз при обследовании нового больного.

### **Список литературы**

1. Вятчин Д.А., Старовойтов В.В., Недзведь А.М. Система принятия решений для диагностики хронических заболеваний кишечника // Сборник докладов X Международной конференции по мягким вычислениям и измерениям SCM'2007, том 2, 2007. – С. 182–185.
2. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. – М.: Горячая линия – Телеком, 2007. – 284 с.
3. Вятчин Д.А. Расстояния между нечеткими множествами типа 2 и их применение к решению задач идентификации // Вестник Военной академии Республики Беларусь, № 3, 2006. – С. 11–17.