

Разработка математических методов формализации профессионального знания врача

Котов Ю. Б.

kotsem@voxnet.ru

Москва, ИПМ РАН

Статья посвящена методам прикладного анализа данных, используемых врачом для принятия решений. Распространен традиционный подход: специалист дает полный набор критериев для исчерпывающего решения задачи, математики вырабатывают процедуру принятия решений по этому набору. В задачах постановки диагноза конкретному больному даже опытный врач обычно использует некоторое количество неотрефлексированных (не осознанных) критериев. Поэтому первая цель математика в таких задачах состоит в выявлении и приведении в строгую логическую форму (формализации) этих критериев. Вторая цель: обеспечить запись и обработку суждений врача в форме, нечувствительной к наличию пробелов в сведениях при сохранении логической корректности на всем протяжении преобразований. Основные требования к критериям (непротиворечивость, эффективность, универсальность и отсутствие синонимов) врач проверить не может. Работа предполагает совместную работу врача и математика.

В процессе профессиональной работы врачу приходится в ряде случаев добавлять новые критерии к рабочему набору (для особо трудных больных), что предполагает использование открытых наборов критериев. Эта проблема носит фундаментальный характер для многих видов познавательной деятельности человека.

Врач в процессе лечения больного часто принимает решения на основе недостаточной информации, когда статистические методы малоэффективны. Математик, работающий с медицинскими данными, вынужден создавать новые, нестатистические методы описания и анализа структур, включающих разнородные данные для моделирования суждений врача в условиях неполной информации.

Логические симптомы. Сведения, используемые врачом в диагностической практике, часто включают тексты, числа и изображения, несопоставимые друг с другом. Врач делает собственные предположения относительно их роли в диагнозе данному больному, которые можно моделировать логическими переменными. Обычная булева логика порождает в задачах диагностики излишне сложные конструкции из-за недоступности некоторых сведений.

Удобной оказалась трехзначная логика [1], разработанная Яном Лукасевичем. Например, одна формула трехзначной логики, использующая

10 логических переменных со значениями (ДА, НЕТ, НЕИЗВЕСТНО) при допустимом количестве неизвестных значений не более 3 заменяет около 800 обычных булевых формул с условиями применимости. Трехзначные логические переменные со значениями (ДА, НЕТ, НЕИЗВЕСТНО) будем называть *логическими симптомами*. К классическому набору операций трехзначной логики была добавлена операция наследования, сохраняющая совпадающие значения. Разработаны программы вычислений с трехзначными логическими переменными в диалоговом режиме.

Объект (больной или отдельное обследование больного) соответствует *симптомному вектору*, каждая координата которого есть логический симптом. Для любых двух симптомных векторов определим меры их близости: *сходство* и *неотличимость* [2], основанные на подсчете совпадающих и неизвестных значений координат. Обе величины достигают 1 у пары полностью совпадающих векторов и равны 0 у полностью несовпадающих. Численные значения этих величин различны при частичном совпадении.

Классы объектов зададим их *масками*, т. е. симптомными векторами с максимальной неотличимостью от объектов «своего» класса и с минимальной — от объектов всех остальных классов. Разработаны алгоритмы генерации масок классов по обучающим выборкам [2].

Задачу классификации объектов удобно решать в несколько этапов. На первом методом *диагностических игр*, предложенным И. М. Гельфандом [3], определяем набор логических симптомов, используемых врачом. Затем, используя данные об известных больных (класс обучения) в пространстве этих симптомов, строим маски классов. И, наконец, используя маски классов, вычисляем неотличимость каждого нового больного от всех масок. Максимальная величина ее указывает на возможную принадлежность больного определенному классу. Простейший метод классификации — пороговый — уже дает удовлетворительные результаты. Ошибки и двусмысленности классификации отдельных больных прямо указывают на недостающие данные или неустойчивые суждения врача, т. е. на содержательные «болевы точки» процесса диагностики. Окончательная классификация проводится по результату вычисления значения критерия.

Проиллюстрируем возможности этого подхода примерами решенных задач диагностики, результаты которых внедрены в практику.

1. После кесарева сечения на теле матки остаются рубцы от хирургических разрезов. Повторные роды несут опасность разрыва матки в случае несостоятельного рубца. Механические свойства миометрия у конкретной пациентки неизвестны и недоступны для исследования. Требовалось дать прогноз опасности для конкретной беременности. Мы прогно-

зировали не механические процессы, а решение опытного врача на трех последовательных этапах (предварительное обследование, решение накануне плановых родов, экстренное решение в родах). Использование технологии логических симптомов позволило резко сократить объем собираемой информации (8 симптомов вместо 450 переменных, названных врачом).

2. Дети, родившиеся живыми от матерей с сахарным диабетом, по-разному адаптируются к самостоятельной жизни. Особенно велика роль раннего периода жизни (первые 7 дней). Для своевременного начала лечения нужен диагноз уже в первые часы жизни. На первом этапе несколько врачей единогласно и независимо указали крайних (тяжелых и благополучных) новорожденных. Неотличимость больных от маски благополучного класса может служить оценкой тяжести новорожденного.

3. При поликлиническом осмотре женщин пожилого возраста для выявления ранних опасных изменений шейки матки используют три метода: ультразвуковое исследование с измерением скорости кровотока методом Доплера, расширенную кольпоскопию и цитологию мазков. Наиболее эффективен (и дорог) первый из них. Задача состояла в оценке эффективности более простых методов. Логические симптомы, связанные с методом кольпоскопии, позволяют решить задачу диагностики. Сравнение индивидуальных наборов логических симптомов каждой больной с двумя масками позволило выделить всех пациентов с неблагоприятным прогнозом.

Итог. Разработан формальный язык логических симптомов, позволяющий моделировать суждения врача, выдвигаемые в процессе постановки диагноза конкретному больному. Диагностическое правило допускает формальную запись (например, в виде ДНФ) достаточно сложных решений. Разработан набор диалоговых программ для обслуживания операций над данными.

Работа выполнена при поддержке РФФИ, проект № 04-01-00434.

Литература

- [1] Карпенко А. С. Логика Лукасевича и простые числа.— М: Наука, 2000. — 319 с.
- [2] Котов Ю. Б. Метод логических симптомов в моделировании профессиональных суждений врача // Информационные технологии и вычислительные системы. — 2005. — № 2. — С. 29-42.
- [3] Гельфанд И. М., Розенфельд Б. И., Шифрин М. А. Очерки о совместной работе математиков и врачей— М: Наука, 1989. — 272 с.