

ПРИНЯТИЯ РЕШЕНИЙ В ЗАДАЧАХ КЛИНИЧЕСКОЙ ДИАГНОСТИКИ В УСЛОВИЯХ НЕОПРЕДЕЛЁННОСТИ

И.Н. Долгополов

Международный центр информационных технологий и систем НАН Украины и МО
e-mail:dolgigo@ukr.net

Основную задачу управления процессами проведения клинической диагностики (КД) можно сформулировать следующим образом: дана совокупность инструментальных, клинических и биохимических средств, в дальнейшем – средств диагностических технологий (СДТ) $T_{\beta, \beta} = 1, \dots, r$, которые в результате проведения исследований $A = \{a\}$, обеспечивают получение первичной информации о состоянии обследуемого пациента. На протяжении диагностической траектории (ДТр), первичная информация, дополняясь, в результате принятия решений (ПР), сменяется вторичной, в виде гипотез о предварительных, правдоподобных диагнозах, затем – конечной, диагностической информацией, на основании которой принимается решение о постановке клинического или достоверного диагноза. Учитывая особенности КД, заключающиеся в необходимости использования как вероятностных, так и детерминированных оценок (предварительный диагноз – вероятностная оценка, клинический или заключительный – достоверная), рассмотрим в качестве универсального распознающего устройства и системы ПР, интеллектуальный конечный автомат Finite-state machine – (ИА). В рамках синтаксического подхода, идентификацию состояния пациента, как основной компоненты технологии постановки диагноза, можно представить следующим образом: известна грамматика W , известно описание объекта в терминах «фразы», например X . Такая постановка задачи позволяет определить принадлежность X к $L(W)$ для объектов с жёстко заданными правилами функционирования. Каждой грамматике можно привести в соответствие некое распознающее устройство в виде автомата. Рассмотрим грамматику W , которой соответствует автомат с конечным множеством состояний. По определению, конечный автомат описывается с помощью пяти характеристик:

$$A_1 = \{Q, l, \delta, q_0, q_f\},$$

где Q – конечное множество состояний; l – множество добавляемых признаков; δ – таблица переходов, содержащая Ql возможных вариантов; q_0 – начальное состояние ($q_0 \in Q$), q_f – множество окончных (финитных, терминальных) состояний, также являющееся подмножеством Q [1].

В зависимости от возможных исходов различают автоматы детерминированный и вероятностный. Если переход из одного состояния в другое не зависит от вероятности возможных исходов и входному воздействию l , состояния q_0 , соответствует следующее состояние q_f . можно говорить не о стохастическом, а детерминированном конечном автомате. Конечный недетерминированный автомат в КД описывается выражением

$$\delta(q_i, i_k) = \{q_3, q_5, \dots, q_1\},$$

которое означает, что каждой паре величин (состояние q_i и входное воздействие i_k) соответствует одно из следующих состояний q_3, q_5, \dots, q_1 . Например, после получения очередного признака, состояние больного можно обобщить в виде предварительного диагноза, который выбирается из некоего множества правдоподобных гипотез с разной степенью вероятности. Или, диагнозу больного конкретного класса состояния (пол, возраст, факторы риска и антириска), с определённой вероятностью соответствуют причины развития заболевания. Детерминированный конечный автомат применяется в разных задачах диагностического цикла, результатом которых может быть только одно решение. В каждой из них, переход имеет следующий вид:

$$\begin{array}{c}
 g(q_i, i_k) = \{q_l\} \\
 \begin{array}{ccc}
 \uparrow & \uparrow & \uparrow \\
 \text{состояние} & \text{входное воздействие} & \text{единственное} \\
 & \text{в данный момент} & \text{следующее} \\
 & & \text{состоянии}
 \end{array}
 \end{array}$$

Например, в задачах постановки диагноза q_i - описывается неким множеством выявленных признаков, i_k - входное воздействие представляет собой недостающий признак $x_{j,j}$, для постановки диагноза D_j , q_l - выставляемый диагноз D_j . По этой же схеме, каждый выбранный признак, необходимый для постановки DS детерминирует *выбор метода определения СДТ* и т.д., согласно схемы оптимизации ДТр.

Для решения задач КД, характеризующихся отсутствием полноты информации используются методы математического программирования, в частности стохастического программирования, включающие решения в условия *риска* и *неопределённости* [2]. Задачи принятия решений в условиях *риска* характеризуются известными вероятностями возможных исходов принимаемых решений и могут быть приведены к возможностям детерминированного конечного автомата. Для решения задач в условиях *неопределённости* вероятности возможных исходов принимаемых решений неизвестны. Например, пусть имеется m допустимых решений, n возможных исходов этих решений и c_{ij} -затраты, связанные с решением i , $i = \bar{1}, m$, при исходе j , $j = \bar{1}, n$. Истинный исход при принятии решений i не известен. Задача заключается в выборе такого допустимого решения i , которое в смысле некоторого критерия является наилучшим. Если бы выбранное решение однозначно приводило к определённому исходу, то задача была бы эквивалентна задачам линейного или нелинейного программирования, так как необходимо выбрать такое решение i при заданном исходе j , для которого величина c_{ij} была бы минимальной. Но исход j не задан, поэтому задачу выбора решения усложняется. При решении задач в условиях риска известны вероятности исходов p_1, \dots, p_n , причём

$$\sum_{j=1}^n p_j = 1.$$

Часто критерием выбора решения i является минимизация средних затрат

$$f_i = \sum_{j=1}^n c_{ij} p_j, \quad i = \bar{1}, m$$

или максимизация вероятности, где c_{ij} - некоторые заданные затраты.

$$f_i = P\{c_{1j} \leq c_{ij} = 1, n\}, \quad i = \bar{1}, m,$$

При решении задач в условиях *неопределённости*, т.е, в условиях когда вероятности исходов $p_j, j = \bar{1}, n$, неизвестны, может быть принят один из следующих критериев оптимальности решения:

1. *Минимизация максимально возможного убытка* (так называемый, гарантированный уровень). Оптимальным является то решением, при котором максимальный убыток достигает минимума.

2. *Критерий минимаксного риска* (критерий Севиджа).

3. *Критерий оптимальности по Парето*.

4. *Принцип недостаточного основания Бернулли* (принцип равновероятности исходов). Если какой-либо исход не вероятней другого, то все исходы принимаются равновероятными. В таком случае задача в условиях неопределённости сводится к задаче в условиях риска. Однако эти методы применимы только для единично/локальных процессов ПР. Для систем, с динамическим, последовательным рядом процессов ПР, необходимы другие методы [3].

Для постановки достоверного диагноза, необходимо разработать алгоритмы построения плана оптимальной ДТр, в соответствии, с которой в СУП КД планировались бы циклы *последовательного и целенаправленного* определения недостающих признаков x_j , $j=1, \dots, N$ средствами ДТ для перевода множества правдоподобных диагнозов в единственно/достоверный. На протяжении ДТр, цикл может неоднократно повторяться, в соответствии с необходимостью отработки очередной гипотезы о возможном диагнозе, до полного узнавания состояния пациента в объёме конечной цели КД - клинического диагноза [4]. Особенностью применения методов линейного программирования в КД является их способность решать задачи динамического характера, в которых результат предыдущего периода состояния определяет исходную информацию последующего. Так например, традиционно в технических системах распознавания считалось, что диагноз, как решение о принадлежности распознаваемого объекта ω соответствующему классу Ω_i , $i=1, \dots, m$, принимается только после определения всей совокупности признаков этого объекта x_1, \dots, x_n . Однако в условиях автоматизации процессов КД возможен другой подход. Задача распознавания должна решаться на основании информации об измеренных к текущему моменту КД каждого очередного признака состояния обследуемого пациента: x_1, x_2, x_3 и т.д. При этом процесс набора дополнительной информации путем измерения последующего признака состояния или прекращение этого процесса происходит в зависимости от результатов полученного решения на каждом этапе исследований. Такая последовательная процедура решения задачи распознавания обязана своим возникновением одному из разделов статистики - последовательному анализу [5]. Применение целенаправленных технологий последовательного и многократного ПР в задачах распознавания, с использованием на каждом этапе исследований возрастающего числа измеренных признаков актуально для клинической медицины, где в существующей практике преобладает, в основном, «веерное» расширение объёмов исследований. Особенно, когда процесс КД протекает в условиях использования многочисленных СДТ, и связан с определенным риском, неопределённостью и целым рядом ограничений Γ на проведение процесса клинических исследований. При этом стоимость выявляемого класса состояния, в объёме достоверного диагноза, не должна превышать установленной величины. Значение расходов U_ω , дальнейшей ДТр по формированию признакового пространства приоритетной гипотезы (ΔI) , усредненное по возможным ее вариантам, определяется правилом R , в соответствии с которым планируется последовательность и объём диагностических мероприятий, т.е. $\bar{U}_\omega = \bar{U}_\omega(R)$. При этом, каждое правило R должно строиться с учетом ограничений Γ , накладываемых на конкретный план мероприятий проводимых исследований. В результате нескольких итераций инструментальных циклов исследований, полученная информация используется для принятия решения о принадлежности состояния диагностируемого пациента к одному из возможных классов состояний, в объёме заключительного диагноза.

Литература

1. Фор А. Восприятие и распознавание образов. – М: Машиностроение, -1989, 272 с.
2. Благодатских В.И. Введение в оптимальное управление (линейная теория). М.: «Высшая школа», 2001, 240 с.
3. Чикрий А.А Конфликтно управляемые процессы. Киев: Наукова думка, 1992, 381 с.
4. Долгополов И.Н. Интеллектуальная система управления процессами клинической диагностики//Проблемы информатики и управления. -2007.- №3.- С. 146-154.
5. Фу К. Последовательные методы в распознавании образов и обучении ЭВМ - М: Наука, 1971, 256 с.