

Н.И. Косарев,

доктор физико-математических наук,
доцент
Сибирский юридический институт
ФСКН России (г. Красноярск)

МОДЕЛЬ БАЗЫ ЗНАНИЙ ЭКСПЕРТНОЙ СИСТЕМЫ, ИСПОЛЬЗУЮЩАЯ ТЕОРЕМУ ГИПОТЕЗ ДЛЯ ОПРЕДЕЛЕНИЯ ВИДА НАРКОТИКА ПО ВНЕШНИМ ПРИЗНАКАМ ЛИЦА, ЕГО УПОТРЕБИВШЕГО

Экспертная система (ЭС) представляет собой программно-аппаратный комплекс, использующий знания эксперта при решении задач в неформализованных областях, не поддающихся алгоритмическому моделированию. Главное ее отличие от обычных программ заключается в том, что последние обрабатывают данные, а ЭС манипулируют знаниями. На самом деле, любой начинающий пользователь персонального компьютера, знакомый с каким-либо языком программирования высокого уровня, представляет себе способ как описания программы, так и ввода в нее начальных данных для решения задачи. С другой стороны, даже квалифицированному программисту трудно представить, каким же образом осуществить ввод знаний в компьютерную программу. Поэтому не случайно ЭС представляет собой интеллектуальную программу, являясь разновидностью системы искусственного интеллекта. Главная причина ее интеллектуальности кроется в базах знаний для хранения тех самых долгосрочных знаний, которые и использует ЭС для решения задач. Ключевым при создании баз знаний является вопрос относительно способа формализации и организации знаний в ЭС.

В теории построения систем искусственного интеллекта существуют несколько моделей организации знаний, анализ которых предполагает достаточно серьезное знание основ математической логики, машинного интеллекта, робототехники и других разделов кибернетики. Для упрощенного понимания того, как в компьютерной программе можно хранить долгосрочные знания, предлагается использовать классическую теорию вероятностей.¹ Здесь речь идет о теореме гипотез, называемой еще формулой Байеса, которая является следствием теоремы умножения и формулы полной вероятности.² Представим себе следующую задачу. Имеется полная груп-

па несовместных гипотез $B_1, B_2 \dots B_n$. Вероятности этих гипотез до опыта известны и равны соответственно $P(B_1), P(B_2) \dots P(B_n)$. Произведен опыт, в результате которого наблюдалось появление некоторого события A . Спрашивается, как следует изменить вероятности гипотез в связи с появлением этого события? Здесь требуется найти условную вероятность для каждой гипотезы. Из теоремы умножения имеем:

$$P(AB_i) = P(A)P(B_i|A) = P(B_i)P(A|B_i), \quad (i=1, 2 \dots n).$$

Если отбросить левую часть предыдущего уравнения, получим $P(A)P(B_i|A) = P(B_i)P(A|B_i), \quad (i=1, 2 \dots n)$.

Из последнего равенства имеем

$$P(B_i | A) = \frac{P(B_i)P(A|B_i)}{P(A)}, \quad (i=1, 2 \dots n).$$

Выражая $P(A)$ с помощью формулы полной вероятности, имеем:

$$P(B_i | A) = \frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \dots + P(B_n)P(A|B_n)}, \quad (i=1, 2 \dots n).$$

Последняя формула и носит название формулы Байеса, или теоремы гипотез.

Теперь рассмотрим прикладную задачу использования теоремы гипотез для создания базы знаний экспертной системы. Пусть в городе N дезоморфин является вторым по популярности наркотиком после героина. После применения героина главными признаками у наркомана являются наличие *маленьких «точечных» зрачков* и *сонливое состояние*. Известно, что после применения дезоморфина появляются такие признаки, как *сужение зрачков совместно с красными глазами*. Еще одним признаком использования дезоморфина является *нарушение сна* у принявшего его наркомана. Предположим, в наркодиспансер поступает пациент со следами передозировки, которому необходимо срочно оказать первую медицинскую помощь во избежание летального исхода. Способ такой помощи зависит от того, что употребил пациент. По внешним признакам необходимо определить вероятность того события, что пациент употребил *героин*. Будем считать, что существует априорная вероятность того, что некоторая часть наркозависимой молодежи употребляют героин, а другая часть – дезоморфин. Обозначим первое событие буквой Γ , а второе – Δ . Предположим, что произошло событие Γ : пациент употребил героин. Пусть априорная вероятность (известная до опыта) события Γ равна $P(\Gamma) = 0,6$, а

априорная вероятность события Д равна $P(D) = 0.3$. Это значит, что из 10 наркозависимых подростков 6 употребляют героин, а 3 – дезоморфин. Априорные вероятности находятся из анализа статистических данных. Вероятности событий Г и Д в общем случае не образуют полную группу, поскольку имеются еще другие виды наркотических средств. На вероятность их применения остается величина 0.1. Признак употребления героина наличие у пациента *маленьких «точечных» зрачков* обозначим Г1. Другой признак, *сонливое состояние*, обозначим Г2. Требуется определить условную вероятность события Г1 при условии, что произошло событие Г (употребление героина). Эту условную вероятность следует определить по поведению пациента при его осмотре наркологом. В данном случае нарколог выступает в роли эксперта. Следует отметить, что признак Г1 не всегда является абсолютно достоверным, то есть эксперт не всегда абсолютно точно может определить наличие данного признака у пациента, но может лишь указать примерно вероятность проявления данного признака. Поэтому условную вероятность события Г1 при условии события Г не следует приравнивать к единице. Этот вопрос следует исследовать статистическим испытанием. Например, если из 100 пациентов, употребивших героин, 85 имеют маленькие «точечные» зрачки, то условная вероятность события Г при условии наступления события Г1 равна $P(G|G1) = 0.85$. Теперь предположим, что сонливое состояние наблюдается примерно у 95 пациентов из 100 принявших героин. Тогда условная вероятность события Г при условии Г2 равна $P(G|G2) = 0.95$. Также нужно учесть, что после употребления дезоморфина у пациента наблюдаются узкие зрачки. Но этот признак проявляется совместно с красными глазами. Поэтому не всегда можно точно определить по зрачкам, употреблял пациент дезоморфин или другой наркотик. Поэтому определим условную вероятность наличия у пациента маленьких «точечных» зрачков, если он употребил дезоморфин, вероятностью, близкой к нулю, например $P(G1|D) = 0.15$. Кроме того, у нас есть еще второй признак – сонливое состояние, обозначенный нами событием Г2, проявляемый у наркозависимых людей, принимающих дезоморфин. Рассуждая приведенным выше способом, введем условную вероятность наблюдения у пострадавшего сонливого состояния при условии употребления им героина. Пусть $P(G2|G) = 0.9$. Таким же путем определим вероятность сонливого состояния при условии употреб-

ления пациентом дезоморфина. Поскольку в последнем случае должно наблюдаться нарушение сна, то положим $P(\Gamma_2|D) = 0.1$.

Теперь рассчитываем вероятность гипотезы о том, что пациент употребил героин. Используя формулу Байеса, имеем:

$$P(\Gamma|\Gamma_1) = \frac{P(\Gamma) \cdot P(\Gamma_1|\Gamma)}{P(\Gamma) \cdot P(\Gamma_1|\Gamma) + P(D) \cdot P(\Gamma_1|D)} = (1)$$
$$\frac{0.6 \cdot 0.85}{0.6 \cdot 0.95 + 0.3 \cdot 0.15} = 0.92$$

$$P(\Gamma|\Gamma_2) = \frac{P(\Gamma) \cdot P(\Gamma_2|\Gamma)}{P(\Gamma) \cdot P(\Gamma_2|\Gamma) + P(D) \cdot P(\Gamma_2|D)} = (2)$$
$$\frac{0.6 \cdot 0.8}{0.6 \cdot 0.9 + 0.3 \cdot 0.1} = 0.947$$

Преыдущими формулами мы рассчитали вероятности двух гипотез употребления пациентом героина при условии появления симптомов Γ_1 и Γ_2 . Данные симптомы могут проявиться одновременно и независимо друг от друга. Следовательно, мы должны рассмотреть сумму двух событий $C = \Gamma_1 + \Gamma_2$, состоящую по определению в появлении хотя бы одного из событий – события Γ_1 , либо события Γ_2 , либо одновременно событий Γ_1 и Γ_2 . Поэтому для расчета полной вероятности такого события C нужно воспользоваться формулой:

$P(C) = P(\Gamma|\Gamma_1) + P(\Gamma|\Gamma_2) - P(\Gamma|\Gamma_1) \cdot P(\Gamma|\Gamma_2)$. Подставив в нее значения, найденные по формулам (1) и (2), получим для $P(C) = 0.99$. Это означает, что если у пациента наркологом обнаружены маленькие «точечные» зрачки и сонливое состояние, то с вероятностью, близкой к единице (точнее, равной 0.99), этот пациент употребил героин.

Возникают закономерные вопросы: где же в такой модели определена роль эксперта в предметной области знаний и как эти знания формализовать и заложить в ЭС? Во-первых, эксперт (в данном случае нарколог) на основе изучения статистического материала должен определить априорные вероятности (известные до опыта) использования того или иного наркотика – вероятности $P(\Gamma) = 0.6$ и $P(D) = 0.3$. Во-вторых, эксперт должен знать наиболее важные признаки, которые являются следствием употребления данного вида наркотического средства. Последние, так же как и априорные вероятности, в виде значений конкретных коэффициентов и вводятся в базу знаний.

В заключение следует отметить, что, используя теорему гипотез, действительно можно построить экспертную систему, кото-

рая кроме теории вероятностей не требует знания специфических разделов прикладной математики и кибернетики. С одной стороны, она уступает современным системам искусственного интеллекта, а с другой – значительно упрощает понимание такого важного с прикладной точки зрения программного продукта, каковым является экспертная система.

¹ Нейлор К. Как построить свою экспертную систему. М., 1991.

² Савельев Л.Я. Элементарная теория вероятностей, 1 : учебное пособие. Новосибирск, 2005.

М.Ю. Селиванова,

кандидат филологических наук,
Управление ФСКН России
по Кировской области

КОНТАКТОУСТАНАВЛИВАЮЩИЕ ЭЛЕМЕНТЫ КАК ОДИН ИЗ ИДЕНТИФИКАЦИОННЫХ ПРИЗНАКОВ В ФОНОСКОПИЧЕСКОЙ ЭКСПЕРТИЗЕ

Фоноскопическая экспертиза предполагает комплексный, всесторонний анализ устной речи говорящего, включающий выявление особенностей его речи на всех уровнях языка. Безусловно, вывод о принадлежности реплик с разных фонограмм одному лицу делается после анализа совокупности всех выявленных лингвистических признаков, определенных на всех элементах, составляющих структуру речи. Мы рассмотрим один лингвистический признак – употребление средств речевого контакта.

Под «речевым контактом» мы имеем в виду установление, укрепление, поддержание связей и отношений с помощью определенного набора вербальных средств, под контактоустанавливающими средствами (средствами речевого контакта) – средства языка, способствующие поддержанию контакта между коммуникантами.

Специфика телефонного разговора заключается в отсутствии визуального контакта между собеседниками. Следовательно, при анализе реплик участников телефонных переговоров можно говорить лишь о вербальных средствах установления и поддержания речевого контакта.