

О ПОДХОДАХ К ИНЖЕНЕРИИ ЗНАНИЙ В СИСТЕМАХ С НЕДЕДУКТИВНЫМ ВЫВОДОМ

В.А. Петрухин, Ю.Н. Манойло

Институт кибернетики им. В.М. Глушкова НАН Украины
03680, Киев, проспект Академика Глушкова, 40,
тел.: (044) 526 3579; e-mail: vapetr@icyb.kiev.ua, yurik@saybervizhn.ru

Статья посвящена вопросам инженерии знаний экспертов при создании и интерпретации математических моделей сложных динамических систем. Рассмотрены подходы к интеграции локальных знаний экспертов о причинно-следственных, логико-функциональных связях параметров изучаемой системы.

The article is dedicated to questions of experts knowledge engineering at creation and interpretation of mathematical models of complex dynamic systems. Approaches of integration of experts local knowledge about cause-effect, logic-functional relations of studied system parameters are considered.

Введение

Использование вычислительной техники и математических методов в медицине – одно из бурно развивающихся направлений информатики. Математические модели в медицине являются моделями нового типа. Создание их представляет собой большое искусство. Такие модели помогают не только решить с помощью современной компьютерной техники сложные многопараметрические задачи диагностики, но и выбрать оптимальные пути лечения. Особый интерес представляет проблема разработки и исследования динамических моделей в медицине, открывающих возможности решения задач не только диагностики состояния больного, но и прогнозирования развития течения патологического состояния организма с учетом различных вариантов лечения, внешних факторов. Основная трудность состоит в отсутствии точных математических моделей изучаемых процессов – часто в наличии имеется только эмпирический опыт специалистов высокой квалификации, который необходимо научиться использовать для построения математических моделей сложных динамических систем. Описания динамических систем – описания последовательной смены явлений, состояний (закон эволюции) – могут быть разнообразными: с помощью аппарата алгебраических, дифференциальных уравнений, дискретных преобразований, графовых представлений, марковских процессов и т.д. Если динамическая система задана уравнением, то постулируется, что каждому состоянию $x(t_0)$ в фазовом пространстве ставится в соответствие состояние $x(t)$, $t > t_0$, куда за время $t - t_0$ переместится фазовая точка, движущаяся по уравнению математической модели системы.

В операторной форме математическую модель можно записать в виде $x(t) = T_t x(t_0)$, где T_t – закон (оператор) эволюции. Если этот оператор применить к начальному состоянию $x(t_0)$, то получим $x(t)$ при $t > t_0$. Для медицины характерно использование качественных параметров и зависимостей между ними. Исследуем подходы к заданию указанного оператора T на основе комплексного анализа информации медицинских баз, использования знаний специалистов-экспертов. Методологическая основа моделирования патологических процессов и состояний человека – системный анализ, центральной процедурой которого является построение обобщенной (единой) модели объекта, отражающей важнейшие факторы и взаимосвязи реальной системы. Исследуемая система – организм человека – имеет сложную внутреннюю структуру, в составе которой могут быть декомпозированы различные подсистемы, представляющие внутренние органы и системы человека, процессы медикаментозного воздействия в процессе лечения, факторы внешней среды, образа жизни и т.д. [1, 2]. Системный подход к моделированию в медицине и биологии был предложен и исследован ранее [3]. Он учитывает то, что организм человека, как объект исследования, представляет собой сложную динамическую систему и характеризуется:

- слабостью теоретических знаний, отсутствием единой количественной, математизированной теории (медицина является ярким примером недедуктивной системы знаний);
- качественным характером знаний о системе, большой долей экспертных знаний при описании, структуризации объекта моделирования, поскольку задачи управления патологическими состояниями человека являются слабоструктурированными;
- высоким уровнем неопределенности исходной информации (внутренней неопределенностью, зависящей от совокупности тех факторов, которые не контролируются лицом, принимающим решение полностью, но он может оказывать на них влияние, а также внешней неопределенностью, обусловленной характером взаимодействия с внешней средой, факторами, которые находятся под слабым контролем врача).

Использование системного подхода к моделированию в медицине и моделей обобщенных динамических систем актуально в ситуациях, когда сложность изучаемых систем такова, что один специалист не в со-

стоянии охватить все взаимосвязи процессов и явлений. Другая отличительная особенность таких систем – отсутствие количественных законов развития процессов изучаемых объектов, так как известны лишь некоторые эмпирические закономерности в узких предметных областях, которыми владеют эксперты-специалисты.

Язык формализации опыта эксперта

Язык формализации экспертного знания является неотъемлемой частью системы представления и интерпретации знаний в динамических предметных средах, использующей процедуры недедуктивного, статистического вывода. Один из самых распространенных формальных языков представления знаний – язык правил типа "если-то", называемых также продукциями. Каждое такое правило есть, вообще говоря, некоторое условное утверждение.

Например:

- если предварительное условие P, то заключение (вывод) C;
- если ситуация S, то действие A;
- если выполнены условия C1 и C2, то не выполнено условие C.

"Если-то" – правила обычно оказываются весьма естественным выразительным средством представления знаний. Они обладают следующими свойствами:

- модульность – каждое правило описывает небольшой, относительно независимый фрагмент, модель знаний;
- возможность постоянной актуализации базы знаний, можно добавлять новые модули правил в базу знаний относительно независимо от других правил;
- удобство модификации (как следствие модульности) – старые правила можно изменять и заменять на новые независимо от других;
- применение вышеперечисленных правил способствует прозрачности системы.

Под прозрачностью понимается способность системы к объяснению принятых решений и полученных результатов. "Если-то" продукции применяют для определения логических отношений между понятиями-параметрами предметной области. Про чисто логические отношения можно сказать, что они принадлежат к "категорическим знаниям", потому что соответствующие утверждения всегда абсолютно верны. Однако в некоторых предметных областях, таких, как медицинская диагностика, преобладают "мягкие" или вероятностные знания. Эти знания являются "мягкими" в том смысле, что говорить об их применимости к любым практическим ситуациям можно только до некоторой степени ("часто, но не всегда"). В таких случаях используют модифицированные "если-то" – правила, дополняя их логическую интерпретацию вероятностной оценкой. Например:

если <условие A> то <заключение B> с уверенностью F.

Вообще говоря, достигнуть определенного понимания предметной области после общения с экспертами и чтения литературы, а затем облечь это понимание в форму представления знаний в рамках выбранного формального языка – искусство, называемое инженерией знаний. Как правило, это сложная задача, требующая значительных усилий. Данная задача требует разработки специальных методик, что особенно важно при исследовании динамики развития патологических процессов в медицине. Правила, содержащиеся в базе знаний, в обобщенном виде имеют вид

$\langle \text{ИмяПравила} \rangle : \langle \text{Условие} \rangle \text{ то } \langle \text{Заключение} \rangle,$

где $\langle \text{Заключение} \rangle$ – это простое утверждение, а $\langle \text{Условие} \rangle$ – набор простых утверждений, соединенных между собой операторами и/или, не.

Синтаксис описания параметра

При описании синтаксиса языка представления формализованного опыта специалистов используется форма Бэкуса – Наура, а также фигурные скобки, в них заключаются повторяющиеся элементы, которые могут встречаться нуль или более раз. Введение любого параметра предполагает задание прежде всего его имени – идентификатора (лексической единицы, используемой экспертами в своей практической работе – полное имя параметра). На практике каждым экспертом для удобства работы могут использоваться короткие имена параметров, имеющие однозначное соответствие уникальному полному имени параметра. Кроме имени параметра задается область изменения значений. Для числовых параметров она задается фиксированием значений границ изменения параметра. Область изменения качественного параметра задается перечислением возможных градаций его значений. При описании параметра может быть введена некоторая базовая размерность значений для того, чтобы эксперты могли пользоваться привычными для них единицами размерности. В конкретном представлении зависимостей параметра данная единица размерности может быть экспертом переопределена (только в рамках представления данного его опыта в записи экспертного знания).

При задании параметра может также быть введено его некоторое нормальное значение, которое принимается (по умолчанию) в качестве начального его значения в процессе проведения экспериментов с моделью (если, конечно, начальное значение, точнее, предыстория динамики параметра не задана специально).

В описании параметра используются числа двух классов: целочисленные и вещественные. Тип числа в описании явно не задается, однако он присутствует в словаре-справочнике параметров изучаемой системы (т.е.

во внутреннем представлении словаря-справочника параметров и используется при переводе элементов языка формализации опыта экспертов во внутреннее представление системы представления и интерпретации знаний в динамических предметных средах). Тип параметра: числовой (целочисленный или вещественный), качественный. Он задается форматом представления чисел, используемых при описании параметра – области изменения его значений, размерности и нормального значения параметра:

<описание параметра> ::= <полное имя параметра><описание области изменения значений параметра> <описание размерности значений параметра><описание нормального значения>

<полное имя параметра> ::= <идентификатор>

<описание области изменения значений параметра> ::= < описание области изменения значений числового параметра>|<описание области изменения значений качественного параметра>

< описание области изменения значений числового параметра>::= (<число> : <число>)

<описание области изменения значений качественного параметра> ::= (< элемент области изменения значений качественного параметра>:{<элемент области изменения значений качественного параметра>})

<нормальное значение параметра> ::= <элемент области изменения значений параметра>|<элемент области изменения значений параметра>-<элемент области изменения значений параметра>

Примеры описания параметров:

ТЕМПЕРАТУРА (34:43), 'РАЗМЕРНОСТЬ': (ГРАДУС), 'НОРМА':(36.5-36.9);
 ПУЛЬСАЦИЯ СОННОЙ АРТЕРИИ (НЕТ:ЛЕГКАЯ:ЗНАЧИТЕЛЬНАЯ: РЕЗКАЯ), 'НОРМА': (ЛЕГКАЯ);
 АРТЕРИАЛЬНОЕ ДАВЛЕНИЕ СИСТОЛИЧЕСКОЕ (0:300), 'РАЗМЕРНОСТЬ' : (ММРТСТ), 'НОРМА':
 (100-130);

Основные положения создания языка представления опыта экспертов сформулированы академиком В.М. Глушковым [1, 2], в котором широко используется принцип умолчания для удобства использования языка в практической работе конечными пользователями. Ставилась задача создания языка максимально близкого к естественному, к тому, которым эксперты привыкли пользоваться в своей повседневной работе. На начальном этапе исследований ставилась задача изучения систем, в которых преобладают качественные параметры и зависимости между ними.

Формализация экспертного опыта (запись на языке формализации опыта эксперта) состоит из вступления, описания параметров, вспомогательных зависимостей и самого высказывания (темпоральной продукции):

<экспертное высказывание>::=<вступление><последовательность описаний>

<высказывание>

<вступление> ::= 'ЭКСПЕРТ'<ФИО эксперта>'ОЦЕНКА'<номер оценки>

<номер оценки> ::= <целое число>

Во вступлении для идентификации оценки (представления опыта) задаются фамилия, имя, отчество эксперта и номер его оценки (их может быть несколько, ибо конкретный специалист может быть экспертом для оценки причинно-следственных связей ряда параметров). В экспертной оценке всегда присутствуют описания параметров, используемых в данной оценке для формулирования высказывания. Кроме того, в экспертной оценке могут описываться временные переменные и вспомогательные зависимости:

<последовательность описаний> ::= <описание>|<описание><последовательность описаний>

<описание> ::= <описание глобальных параметров системы>|<описание темпоральных вспомогательных переменных>|<описание вспомогательных зависимостей> |<описание главного параметра>

В экспертной оценке описываются параметры, необходимые для формулирования взаимосвязей "главного" параметра оценки, причинно-следственная зависимость которого от других параметров выражает данный элемент базы знаний. Описание параметра в экспертной оценке преследует несколько целей. Самое минимальное описание может состоять только из указания полного и сокращенного имени параметра, причем сокращенное имя параметра действительно только внутри данной экспертной оценки. Тогда в экспертной оценке доступ к значениям параметра возможен только через использование его сокращенного имени, что удобнее, и именно

такими сокращениями пользуются специалисты в процессе своей работы. На этом описание параметра в экспертной оценке может быть и закончено. Однако часто необходимо в экспертной оценке использовать единицы размерности, отличные от используемых при описании параметра в словаре-справочнике параметров системы. Эта возможность достигается введением при описании параметра в экспертной оценке новых единиц размерности. При трансляции (переводе во внутреннее представление) будет обеспечено соответствие новых единиц размерности единицам, введенным в процессе параметризации изучаемого объекта исследований.

Кроме того, наличие описания параметров в экспертной оценке дает возможность иметь характеристики параметров в процессе работы, не обращаясь к словарю-справочнику параметров системы. В экспертной оценке можно в некотором смысле переопределить параметр – область изменения значения числового параметра разбить на ряд интервалов значений и внутри данных интервалов не различать значений параметра. Таким образом, числовой параметр становится параметром качественного типа (таким параметрам в процессе перевода во внутреннее представление присваивается смешанный тип значений). Сказанное можно подкрепить следующим примером: часто в различных суждениях используют не точное значение температуры, а разбивают область изменения значений на ряд числовых интервалов:

РЕКТАЛЬНАЯ ТЕМПЕРАТУРА (<36.5 : 36.5 – 36.9 : 36.9 – 38 : >38).

Различные специалисты могут пользоваться в своих суждениях различными разбиениями области изменения значений параметра. Однако в оценке нельзя изменять имя параметра (следует придерживаться имени параметра из словаря-справочника), для числовых параметров нельзя изменять границы интервалов области изменения значений. Нельзя изменять структуру описания размерности (количество единиц размерности в «числителе» и «знаменателе» описания размерности, а если единица размерности не является единицей размерности времени, массы и длины – изменять идентификатор, используемый при описании размерности):

<описание глобальных параметров системы> ::= 'ПАРАМЕТРЫ':<описание параметров>

<описание параметров> ::= <описание параметра>; <описание параметра>; | <описание параметра>; { | <описание параметра>; }

<описание параметра> ::= <сокращенное имя параметра> - <полное имя параметра><область изменения значений параметра> <описание размерности параметра>

<сокращенное имя параметра> ::= <идентификатор>

<полное имя параметра> ::= <идентификатор>

<область изменения значений параметра> ::= ; | <область изменения значений качественного параметра> | <область изменения значений числового параметра> | <область изменения значений параметра смешанного типа>

<область изменения значений параметра смешанного типа> ::= (< первый элемент описания области изменения значений значений параметра смешанного типа> : { <число> - <число>: } <последний элемент описания области изменения значений параметра смешанного типа>

< первый элемент описания области изменения значений параметра смешанного типа> ::= <число> | <число>

<последний элемент описания области изменения значений параметра смешанного типа> ::= <число> | > <число>

<описание главного параметра> ::= 'ГЛАВНЫЙ' – <имя параметра>

При формулировании экспертной оценки по умолчанию предполагается, что в качестве темпоральной переменной используется идентификатор *t*. Поэтому чаще всего описание темпоральных переменных отсутствует.

С целью удобства записи экспертного высказывания в экспертной оценке могут быть описаны некоторые вспомогательные зависимости:

<описание вспомогательных зависимостей> ::= <пусто> | 'ПУСТЬ': <описание зависимости>; { <описание зависимости>; }

<описание зависимости> ::= <описание функции> | <описание системы алгебраических уравнений> | <описание системы дифференциальных уравнений>

<описание функции> ::= <заголовок> = <выражение>

<заголовок> ::= <имя функции> | <имя функции> (<свободная темпоральная переменная>)

<имя функции> ::= <идентификатор>

При описании вспомогательных зависимостей в реальной версии системы используется только описание функции. При дальнейшем развитии систем предполагается введение средств описания систем алгебраических, дифференциальных уравнений и др.

В экспертном высказывании заключена сущность экспертной оценки. Экспертное высказывание может содержать несколько альтернатив с их степенями уверенности:

<высказывание> ::= 'ВЫСКАЗЫВАНИЕ': <номер варианта> 'С ВЕР' (<вероятность>) <элементарное высказывание>; { <номер варианта> 'С ВЕР' (<вероятность>) <элементарное высказывание>; }

<номер варианта> ::= < латинское число>

<элементарное высказывание> ::= 'ЕСЛИ' <ситуация> 'ТО' <высказывание1> <высказывание2>

<высказывание1> ::= <высказывание3> | <высказывание3>; <высказывание1 >

<высказывание2> ::= <пусто> | 'ИНАЧЕ' <высказывание1 >

<высказывание3> ::= 'С ВЕР' (<вероятность>) <высказывание4> | <высказывание3> <знак логической операции> <высказывание 3>

<высказывание 4> ::= <сокращенное имя главного параметра> (<темпоральное выражение >) = <выражение20> <выражение21> <знак отношения 2 > <сокращенное имя главного параметра> (<темпоральное выражение>) <знак отношения2> <выражение21>

<выражение20> ::= <константа0> | <сокращенное имя параметра> (<темпоральное выражение>) | <имя функции> | <имя функции> (<темпоральное выражение >) | <связанная темпоральная переменная> | <выражение20> <знак арифметической операции > <выражение20>

<вероятность> ::= <число>

<ситуация> ::= <элементарная ситуация> | <ситуация> <знак логической операции> <ситуация>

<элементарная ситуация> ::= < ситуация0> | <ситуация 1>

< ситуация0> ::= <выражение20> <знак отношения> <выражение20>

< ситуация1> ::= <сокращенное имя параметра> (<темпоральное выражение4>) <знак отношения> <константа1>

<темпоральное выражение> ::= <темпоральное выражение0> | <темпоральное выражение2>

<темпоральное выражение4 > ::= <темпоральное выражение1> | <темпоральное выражение3>

<темпоральное выражение3 > ::= <темпоральное выражение2 > : <темпоральное выражение2>

Приведем пример экспертного высказывания.

'ВЫСКАЗЫВАНИЕ' : I. 'С ВЕР' (0.8) 'ЕСЛИ' $Y(t - 1 \text{ ГОД}) \geq 10$ 'ТО' $1 \leq X(t) \leq 5$;

II. 'С ВЕР' (0.2) 'ЕСЛИ' $Y(t - 1 \text{ ГОД}) \geq Y2(t - 0.5 \text{ ГОД})$ 'ТО' $X(t) = 1$;

В данном примере при выполнении ситуации первой альтернативы главный параметр $X(t)$ может принять любое значение в интервале от 1 до 5.

Особенностью языковых средств является возможность при формулировании высказывания для «главного параметра» достаточно просто определять интервал области изменения значений данного параметра (т.е. иметь неточные суждения о его возможных значениях в качественной оценке экспертом ситуации).

Эксперименты психологов показывают, что максимальное количество структурных единиц информации, за которыми человеческий мозг может одновременно следить, приблизительно равно семи плюс-минус два. Вероятно, это связано с объемом краткосрочной памяти у человека. Дополнительным ограничивающим фактором является скорость обработки мозгом поступающей информации: на восприятие каждой новой единицы информации ему требуется около 5 секунд. Таким образом, исследователи оказались перед серьезной дилеммой. Сложность динамических систем возрастает, но способность нашего мозга справиться с этой сложностью ограничена. Каков же выход из этого создававшегося затруднительного положения? Отметим, что способ преодоления данного информационного барьера в управлении и моделировании сложными системами был извес-

тен еще в древности – divide et impera (разделяй и властвуй), на него обращал внимание академик В.М.Глушков при оценке сложности систем управления экономикой государства [1]. При разработке математической модели сложной динамической системы необходимо разделять ее на все меньшие и меньшие подсистемы, каждую из которых можно исследовать и изучать законы ее развития независимо (именно для этих целей и привлекается опыт экспертов-специалистов). В этом случае не превышает пропускная способность человеческого мозга – для понимания любого уровня системы необходимо одновременно держать в уме информацию лишь о немногих ее частях (относительно не обо всех). В самом деле, декомпозиция вызвана сложностью изучаемого объекта, поскольку именно сложность вынуждает делить пространство его состояний [1,2]. Необходимость формулирования логико-функциональных связей в динамических предметных средах задачу существенно усложняет. Поэтому актуальной является задача построения автоматизированных процедур извлечения знаний из динамических баз данных, хранящих реализации «траекторий» динамической системы, пользуясь которыми в автоматизированном режиме специалист может формализовать свой опыт при построении причинно-следственных, логико-функциональных связей параметров изучаемого сложного объекта.

Интеграция формализованного экспертного знания

Пусть изучаемая система описывается множеством параметров (внутренних, управляющих, параметров внешней среды): $\{ X | x_i(t), i = 1, L \}$. Обобщенной динамической системой (ОДС) назовем совокупность $\{ X, ЭО, A, \delta \}$, где X – множество параметров ОДС; ЭО – множество локальных причинно-следственных, логико-функциональных зависимостей между параметрами ОДС; $A = \|\alpha_{ij}\|$ – матрица весовых коэффициентов экспертных оценок, δ – единица системного времени. $X = ВП \cup УП \cup ПВС$, где $ВП$ – внутренние параметры; $ПВС$ – параметры внешней среды, $УП$ – параметры управления. На множестве внутренних параметров определяется группа параметров, служащих переменными функционала качества управления – множество параметров качества управления (ПКУ). С помощью экспертов устанавливаются зависимости вида $x_i^j(t) = f_i^j(x_{i_1}(t-t_{i_1}), x_{i_2}(t-t_{i_2}), \dots, x_{i_k}(t-t_{i_k}))$, причем $x_i \in ВП, 1 \leq i \leq M; M$ – количество внутренних параметров ОДС, где f_i^j – плотность распределения вероятности значения параметра x_i . Пусть имеется L параметров $\{ x_i | i = 1, L \}$ обобщенной динамической системы, причинно-следственные связи параметра x_i оценивают (формулируют набор эвристических правил) m_i экспертов. Каждый из экспертов на основании своего опыта (или пользуясь некоторыми процедурами извлечения знаний) формулирует локальные причинно-следственные связи этого параметра в динамике, результатом применения которых являются возможные значения параметра с их вероятностью с учетом значения выбранной единицы системного времени δ [4, 5].

Пусть в базе данных имеются «исторические» данные N траекторий – предьстории развития изучаемой системы. Обозначим значение i -го параметра k -й траектории в момент времени t как $\hat{x}_i^k(t)$. Все используемые значения времени t кратны δ . Задача состоит в том, чтобы получить обобщенное мнение всех экспертов, которое давало бы более точный прогноз, чем мнение каждого из них, – построить обобщенную экспертную оценку. Требуется построить обобщенную экспертную оценку $f_{x_i}(t, x) = \sum_{j=1}^{m_i} \alpha_j \cdot f_{x_i}^j(t, x)$, где $\sum_{j=1}^{m_i} \alpha_j = 1$. Задача сводится к поиску значений вектора, удовлетворяющего некоторому критерию оптимальности.

Задача 1. Запишем критерий оптимальности прогнозирования следующим образом:

$$\text{вектор } \alpha_j = \operatorname{argmax} \sum_{k=1}^N f_{x_i}(t, \hat{x}_i^k(t)).$$

$$\text{Подставим значение } f_{x_i}(t, \hat{x}_i^k(t)): \quad \alpha_j = \operatorname{argmax} \sum_{k=1}^N \sum_{j=1}^{m_i} \alpha_j \cdot f_{x_i}^j(t, \hat{x}_i^k(t)).$$

$$\text{Поменяем местами суммы:} \quad \alpha_j = \operatorname{argmax} \sum_{j=1}^{m_i} \sum_{k=1}^N \alpha_j \cdot f_{x_i}^j(t, \hat{x}_i^k(t)).$$

$$\text{Вынесем за внутреннюю сумму } \alpha_j: \quad \alpha_j = \operatorname{argmax} \sum_{j=1}^{m_i} \alpha_j \cdot \sum_{k=1}^N f_{x_i}^j(t, \hat{x}_i^k(t)).$$

Очевидно, что для достижения поставленной цели, необходимо иметь $\alpha_j = 1$ для того j , при котором сумма $\sum_{k=1}^N f_{x_i}^j(t, \hat{x}_i^k(t))$ максимальна, или, другими словами, учитывать мнение только одного эксперта, у которого наилучшие результаты на «исторических» данных.

Задача 2. Рассмотрим другой критерий оптимальности функции прогнозирования $\alpha_j = \operatorname{argmax}_{k=1\dots N} f_{x_i}(t, \hat{x}_i^k(t))$, т.е. наихудший результат прогнозирования должен быть максимально хорошим.

Подставим значение $f_{x_i}(t, \hat{x}_i^k(t))$:

$$\min_{k=1\dots N} \left(\sum_{j=1}^{m_i} \alpha_j \cdot f_{x_i}^j(t, \hat{x}_i^k(t)) \right) \rightarrow \max.$$

Пусть $\beta_k^j = f_{x_i}^j(t, \hat{x}_i^k(t))$.

Преобразуем поставленную задачу оптимизации следующим образом:

$$\min_{k=1\dots N} \left(\sum_{j=1}^{m_i} \alpha_j \cdot \beta_k^j \right) \rightarrow \max.$$

Переформулируем данную задачу:

$$\forall k : \sum_{j=1}^{m_i} \alpha_j \cdot \beta_k^j \geq c, \quad c \rightarrow \max.$$

Вектор градиента функции $\sum_{j=1}^{m_i} \alpha_j \cdot \beta_k^j$ обозначим \vec{g}_k :

$$\vec{g}_k \left(\frac{1}{m_i \cdot \beta_k^1}; \frac{1}{m_i \cdot \beta_k^2}; \dots; \frac{1}{m_i \cdot \beta_k^{m_i}} \right).$$

Обозначим $e_k^j = \frac{1}{m_i \cdot \beta_k^j}$.

Результирующий вектор градиента всех функций $\vec{g} = \frac{\sum_{k=1}^N \vec{g}_k}{N}$,

т.е. $\vec{g}(e^1; e^2; \dots; e^{m_i})$, где $e^j = \frac{\sum_{k=1}^N e_k^j}{N} = \frac{\sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}{N}$.

Координаты точки на плоскости $\sum_{j=1}^{m_i} \alpha_j = 1$ будут равны координатам вектора \vec{g} , умноженного на c , т.е. необходимо решить систему уравнений

$$\vec{g} \cdot c = (\alpha_1; \alpha_2; \dots; \alpha_{m_i}) \quad \text{при условии} \quad \sum_{j=1}^{m_i} \alpha_j = 1.$$

Подставим координаты вектора и выпишем систему уравнений полностью:

$$c \cdot \frac{\sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}{N} = \alpha_j, \quad 1 \leq j \leq m_i, \quad \sum_{j=1}^{m_i} \alpha_j = 1.$$

Из этой системы уравнений находим c и α_j . Количество неизвестных совпадает с количеством уравнений, т.е. можно получить точное решение задачи. Подставим α_j из первого уравнения во второе:

$$\sum_{j=1}^{m_i} c \cdot \frac{\sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}{N} = 1.$$

Откуда следует, что $c = \frac{N}{\sum_{j=1}^{m_i} \sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}$.

Подставив c в первое уравнение, получим требуемые значения вектора α_j :

$$\alpha_j = \frac{N}{\sum_{l=1}^{m_i} \sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^l}} \cdot \frac{\sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}{N} = \frac{\sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^j}}{\sum_{l=1}^{m_i} \sum_{k=1}^N \frac{1}{m_i \cdot \beta_k^l}} = \frac{\sum_{k=1}^N \frac{1}{\beta_k^j}}{\sum_{l=1}^{m_i} \sum_{k=1}^N \frac{1}{\beta_k^l}}.$$

Таким образом, получено аналитическое решение поставленной задачи нахождения коэффициентов линейной «свертки мнений экспертов».

Заключение

Приведенные в статье подходы позволяют решать задачу построения обобщенного экспертного знания на основе интеграции формализованного опыта нескольких специалистов, участвующих в эвристической оценке причинно-следственных связей параметров изучаемой системы. Применение рассмотренных подходов может быть использовано при интерпретации знаний экспертов.

1. Глушков В.М. Основы безбумажной информатики. – М.: Наука, 1982. – 552 с.
2. Глушков В.М., Петрухин В.А., Попов А.А. Системный подход к моделированию в медицине // Кибернетика и вычисл. техника, 1977. – Вып. 36. – С. 3 – 6.
3. Петрухин В.А. О языке формализации опыта экспертов системы представления и интерпретации знаний в динамических предметных средах // Пробл. программирования; 2002. – № 1–2. – С. 441 – 446.
4. Петрухин В.А., Манойло Ю.Н. Средства информационного обеспечения системы автоматизации сбора и обработки данных в комбустологии // Реєстрація, зберігання і обробка даних. – 2003. – № 2. – С. 109 – 119.
5. Петрухин В.А., Манойло Ю.Н. Процедура извлечения знаний из динамических баз данных // Компьютерная математика. – 2003. – № 1. – С. 75–87.
6. Петрухин В.А., Манойло Ю.Н., Черевко О.В. О подходах к интеграции формализованного опыта экспертов // Там же. – 2005. – № 3. – С. 122–127.