

УДК 004.02

## МЕТОД РЕШЕНИЯ ЗАДАЧИ ЗАПРОСА ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

А.С. Клещев, В.В. Грибова, Е.А. Шалфеева

*Институт автоматизации и процессов управления ДВО РАН, Владивосток, Россия  
kleshev@iacp.dvo.ru, gribova@iacp.dvo.ru, shalf@iacp.dvo.ru*

### Аннотация

В статье рассматривается задача сбора необходимой информации о ситуации, обычно решаемая в процессе диагностики для сокращения числа гипотез о диагнозе. На практике при решении этих двух задач важно учитывать все виды причинно-следственных отношений, изменчивость значений признаков с течением времени, влияние событий на динамику значений признаков. Предложенный в статье алгоритм решения этой задачи сводит её к выявлению дифференциаторов среди известных измеряемых признаков. Выявление дифференциаторов производится с использованием знаний о развивающихся внутренних процессах и влиянии на них внешних событий. Такой алгоритм применим для решения задачи запроса дополнительной информации для распознавания ситуаций в различных областях деятельности.

**Ключевые слова:** онтология предметной области, задача диагностики, множество гипотез, дифференцирующий признак, алгоритм решения задачи.

**Цитирование:** Клещев, А.С. Метод решения задачи запроса дополнительной информации / А.С. Клещев, В.В. Грибова, Е.А. Шалфеева // Онтология проектирования. – 2017. – Т. 7, №3(25). – С. 310-322. – DOI: 10.18287/2223-9537-2017-7-3-310-322.

### Введение

Одним из средств повышения эффективности диагностики является использование технологии экспертных систем (ЭС), как средства, позволяющего принять во внимание значительное количество диагностических признаков и минимизировать ошибки, связанные с субъективными факторами. К настоящему времени разработано большое их количество в различных предметных областях (ПрО): диагностика неисправностей в механических и электрических устройствах, медицинская диагностика, диагностика болезней животных, идентификация состояния сельскохозяйственных объектов, диагностика в эпизоотологии и др. Различаются принципы построения диагностических ЭС. Так, некоторые ЭС ориентированы на выдвижение множества полезных гипотез, другие - на выдвижение одной наиболее вероятной гипотезы, в ряде диагностических ЭС используется принцип построения, основанный на конструктивной критике выдвинутой пользователем гипотезы. Наиболее распространены ЭС, которые поддерживают задачу диагностики, предлагая специалисту множество полезных гипотез, которое далее должно быть уменьшено. Однако известные из литературы методы её решения не учитывают одновременно все виды причинно-следственных отношений, промежутки времени, прошедший с момента начала развития диагностируемого процесса, изменчивость значений признаков с течением времени, влияние событий на динамику значений признаков. Целью работы является описание алгоритма запроса дополнительной информации для сокращения множества гипотез о диагнозе, позволяющего снять указанные ограничения.

## 1 Проблемы и обзор решений

### 1.1 Постановки задачи запроса дополнительной информации

При обсуждении спектра задач, решаемых в ПрО на основе экспертных знаний, нередко упоминается необходимость определения того, какой информации не хватает, чтобы получить множество решений. Чаще всего необходимость запроса дополнительной информации возникает при решении задачи диагностики [1-8], реже в других задачах («с неопределёнными знаниями») для доопределения недостающей информации [9]. В ряде ПрО решаются задачи относительно систем или объектов, у которых различные процессы протекают во времени. Тогда под диагностикой понимается распознавание внутреннего процесса, который не присущ диагностируемому объекту при нормальном его функционировании (в медицине - заболевание). Нередко под диагностикой понимается не только распознавание типа отклонения параметров функционирующей системы от нормативных, но и выявление причин, приведших к возникновению этих отклонений в системе.

При решении задачи диагностики специалисты вынуждены одновременно решать и другую задачу – сбора необходимой информации о ситуации, а именно: решать, какие исследования следует выполнить, чтобы на их основе принять верное решение. Среди ЭС, решающих задачу диагностики, можно выделить класс таких, которые помимо решения задачи диагностики дают рекомендации относительно того, какая информация о ситуации необходима для этого [1, 3-8]. Их можно разделить на два подкласса:

- ЭС, в которых задача запроса информации не отделяется от задачи диагностики, а решается в ходе выполнения общего для обеих задач алгоритма [6, 10];
- ЭС, в которых эти две задачи разделены [1, 3, 4, 6-8].

Информация, получаемая при обследовании, используется далее как для выдвижения такого множества гипотез о диагнозе, в котором должна содержаться правильная гипотеза, так и для сокращения этого множества до одной гипотезы. Среди ЭС, в которых задача запроса информации отделяется от задачи диагностики, можно выделить два пересекающихся подкласса:

- ЭС, в которых запрашивается информация для выдвижения гипотез о диагнозе [3];
- ЭС, в которых запрашивается информация для сокращения множества гипотез [1, 4, 6-8].

В ЭС, в которых запрашивается информация для сокращения множества гипотез, иногда рассматривают диагноз, содержащий несколько отклонений от нормы (например, заболевание) [3, 4], и ЭС, в которых диагноз «простой» (содержит один класс отклонений от нормы) [1, 2, 6, 8]. Для последнего случая формальная постановка задачи запроса дополнительной информации [11] состоит в следующем.

**Дано:** База знаний (БЗ, англ. KB), содержащая описание множества диагностических классов, классы обладают свойством делимости;

множество гипотез  $H_{R,KB}$  для некоторого объекта диагностики, такое, что мощность множества  $H_{R,KB}$  не меньше двух;

описание диагностируемой ситуации - набор известных значений признаков  $R$  объекта диагностики (наблюдений).

**Найти:** такой запрос дополнительной информации ( $Q$ ) о ситуации – значение (обычно или желательно одного) признака объекта диагностики, с которым новое множество гипотез будет иметь меньшую мощность, чем множество гипотез  $H_{R,KB}$ .

Будем называть для простоты диагнозом сам диагностируемый процесс (один класс отклонений от нормы). И будем полагать, что БЗ удовлетворяет условию делимости классов, если для любой пары  $\langle R, \text{диагноз} = \text{класс}_i \rangle$ , для которой в БЗ есть истинные предложе-

ния, в этой же БЗ существуют ложные предложения относительно любой пары ( $\langle R, \text{класс} = \text{класс}_j \rangle$ ), при любом  $j \neq i$  (т.е. всегда есть предложение, не позволяющее этой ситуации принадлежать какому-то другому классу).

Как результаты наблюдения, так и запрашиваемая дополнительная информация традиционно относятся к признакам объекта диагностики, представляющим собой значения измеряемых или наблюдаемых (визуально или контактно) свойств объекта и к связанным с объектом внешним событиям.

Суть задачи запроса дополнительной информации для сокращения множества гипотез о диагнозе состоит в следующем: по имеющимся признакам объекта диагностики, для которых существует более одной гипотезы о диагнозе, требуется предложить дополнительный признак, значение которого, полученное в результате наблюдения или измерения, позволит сократить множество гипотез.

## 1.2 Обзор предлагаемых методов

В литературе рассматриваются подходы к решению задачи сокращения множества гипотез о диагнозе, включающие следующие методы запроса дополнительной информации для распознавания:

- запрашиваются значения таких признаков, что какое бы значение ни было получено в результате наблюдения признака, множество гипотез о диагнозе всегда будет сокращено [6];
- запрашиваются значения таких признаков, что при получении некоторых значений такого признака множество гипотез о диагнозе может быть сокращено; при ранжировании таких признаков могут дополнительно учитываться априорные вероятности существования диагнозов-гипотез, а также вероятности появления различных значений этих признаков для них;
- запрашиваются значения нескольких признаков одновременно, если в результате их наблюдения могут быть получены такие комбинации их значений, при которых множество гипотез о диагнозе может быть сокращено, - «синергетические эффекты» [12].

В литературе иногда помимо значений признаков рассматриваются продолжительность присутствия тех или иных значений этих признаков (что решает проблему распознавания внутреннего процесса лишь частично).

Такие методы решения задачи запроса дополнительной информации для распознавания опираются на дополнительную информацию в БЗ.

Например, в БЗ вводятся связи между диагнозом и совокупным признаком (синдромом – в медицине), элементом которого может быть и событие. Иногда для элементов в этой совокупности требуется соблюдение частичного порядка: порядок появления признаков у объекта диагностики, чтобы событие предшествовало внешнему признаку, в том числе с указанным интервалом [12]. Другой пример дополнительной информации в БЗ: вводится классификация диагнозов, так что у классов диагнозов указываются (общие) признаки [3, 5], либо вводится иерархия стадий рассуждения специалиста [10, 12].

Используемые БЗ могут содержать любую совокупность предложений-утверждений, таких как:

- множество признаков с их необходимыми и возможными значениями для каждого диагноза;
- множество признаков с их значениями и события, которые характерны для каждого диагноза;

- множество признаков (с их значениями) с продолжительностью присутствия отдельных признаков или их совокупности;
- упорядоченное множество признаков (например, с указанием вероятности их проявления при данном диагнозе либо с использованием нечётких множеств [13]);
- общие признаки для группы (класса) диагнозов.

## 2 Метод ранжирования дифференциаторов

### 2.1 Постановка задачи запроса дополнительной информации для систем или объектов, у которых внутренние процессы развиваются во времени

Протекающие во времени процессы внутри различных систем или объектов часто связаны между собой посредством причинно-следственных отношений. В этом случае в условиях отношений могут присутствовать внешние факторы, внешние события, происходящие в различные моменты времени [9, 14, 15, 16], постоянные во времени признаки объекта диагностики, иногда называемые индивидуальными характеристиками, в следствиях отношений - признаки (с их значениями), проявляемые объектом диагностики. Например, при диагностике состояния сельскохозяйственных культур в качестве условия может выступать время, прошедшее после обработки растений пестицидами и гербицидами, при диагностике некоторого заболевания - перенесённые травмы и заболевания, приём лекарственных препаратов. Внешними факторами для таких объектов могут быть время года (сезон), температура воздуха, индивидуальные характеристики - возраст (растения, человека), пол.

В этом случае могут учитываться:

- промежуток времени, прошедший с момента начала развития диагностируемого внутреннего процесса, не присущего объекту (системе);
- влияние событий на значения признаков и их изменение во времени;
- влияние индивидуальных характеристик объекта (системы) на проявления этого внутреннего процесса.

Признаки, значения которых могут изменяться с течением времени (на протяжении периода диагностики), иногда называют *темпоральными* или динамическими.

Терминология ( $\Sigma$ ) таких ПрО, где рассматриваются протекающие во времени процессы, связанные между собой посредством причинно-следственных отношений, включает термины, интерпретациями которых являются значения, функциональные соответствия или отношения. Функции и отношения, являющиеся интерпретациями функциональных и предикатных символов, могут зависеть не только от времени (упорядоченное множество моментов которого обозначим  $\{t_0, \dots, t_k\}$ ), но и от координат пространства и других свойств ПрО.

*Онтология ПрО* состоит из  $\Sigma$  и множества предложений (аксиом) на языке  $\Sigma$ , представляющих те свойства терминов онтологии, которые определяются соглашениями [11, 17].

Обозначим распознаваемые внутренние процессы как  $F_{in}$  (множество функциональных символов, зависящих от времени, значения которых - состояния процесса); все признаки (значения которых могут меняться у объектов диагностики) -  $F_{ex}$ ; внешние факторы и события, которые могут иметь место, -  $F_{ev}$ ; признаки, значения которых наблюдались у объекта в разные моменты времени -  $R_{ex}(t_0, \dots, t_k)$ ; наблюдаемые индивидуальные характеристики *объекта* (системы) -  $R_O$ ; факторы и события, наблюдаемые в некоторые моменты времени -  $R_{ev}(t_u, \dots, t_v)$ . Обозначим диагнозы (внутренние процессы, не присущие *объекту*) как  $\Delta \subseteq F_{in}$ . Для причинно-следственного отношения удобно ввести предикатный символ ( $P_c$ ), аргументами которого являются *причины, воздействующие факторы* (необязательны), *следствия и время* (необязательно). Примечание: *воздействующие факторы* - индивидуальные характе-

ристики, а также те внешние факторы или события, которые не являются непосредственной причиной.

Определение онтологических соглашений вышеописанных ПрО осуществляется через функциональные соответствия или отношения вышеуказанных терминов. Постановка задачи запроса дополнительной информации, учитывающая вышеописанные свойства ПрО, такова.

Дано: БЗ, согласованная с онтологией ПрО, удовлетворяющая условию делимости классов-диагнозов; результаты наблюдения ситуации (индивидуальные характеристики, признаки *объекта*, внешние факторы и события)  $R = R_O \cup R_{ex}(t_0, \dots, t_k) \cup R_{ev}(t_u, \dots, t_v)$ , такие, что для них выдвинуто множество гипотез  $H_{R,KB} = \{\Delta_l, \dots, \Delta_m\}$ , имеющее мощность не меньше двух.

Найти: такой запрос дополнительной информации Q об *объекте* (или его ситуации) для результатов R, что для R, объединённого с ответом AQ на этот запрос, новое множество гипотез имеет меньшую мощность, чем множество гипотез  $H_{R,KB}$  (т.е. хотя бы одна причинно-следственная модель  $P_c(R, \Delta_j)$  перестанет быть истинной:  $\neg P_c(R \cup AQ, \Delta_j)$ ).

Здесь в БЗ входят предложения ( $A_\Sigma$ ) о свойствах ситуаций каждого класса  $\Delta_i$  из  $\Delta$ , среди которых в общем случае есть:

- варианты развития внутреннего процесса, связывающие диагноз  $\Delta_i$  с изменениями значений признака и зависящие от периода развития этого процесса (например, для диагноза «пневмококковый конъюнктивит» характерны признаки: покраснение глаз. Присутствие = «имеется». Локализация с вариантами динамики: = «справа, слева, справа И слева постоянно»; = «справа 1-2 суток, затем справа И слева»; = «слева 1-2 суток, затем справа И слева» [14]; для «отказа в навигационной системе» автономного подводного робота (АПР) характерны признаки с вариантами проявления (динамики):
  - несанкционированное выключение питания двигателя (питание двигателя = «выключено») ИЛИ
  - высокое энергопотребление со стороны движительно-рулевого комплекса (ДРК), характерного для модификации имеющегося ДРК, в течение промежутка времени, большего, чем требуется для преодоления препятствия в рамках миссии АПР [15];
- варианты развития (или появления) внутреннего процесса, связывающие значения некоторого события с диагнозом (например, при диагнозе «пневмококковый конъюнктивит» его возможные причины:
  - через 1-2 суток после начала предшествующего заболевания, такого как острая респираторная вирусная инфекция, грипп;
  - через 1-2 суток после загрязняющего воздействия окружающей среды, такого как: «попадание в глаз пыли», «купание в непроточном водоеме» [14];
 при диагнозе для АПР «зацепился за препятствие» характерно поступление совокупности сигналов (от соответствующих датчиков): «ошибка стабилизации скорости», «ошибка стабилизации по вертикальному каналу» и «ошибка стабилизации по горизонтальному каналу», а его Возможные причины – цепочка (последовательность) событий, предшествующая появлению указанных признаков:
  - «зависание» (в течение времени  $t > 5$  сек) на текущей глубине lev1 с заданием максимального заднего хода (в течение времени  $t > 3$  сек);
  - затем - попытка перейти и стабилизироваться на глубине lev2, на 2-3 метра меньшей, чем lev1,
  - затем – движение (переход) на программную высоту) [15];
- варианты развития внутреннего процесса, изменённого воздействием события, связывающие диагноз и событие с некоторым значением признака; например, «острый аппенди-

цит, перфоративная язва: обезболивающие таблетки приводят к временному устранению боли (что искажает клиническую картину)»);

- варианты развития внутреннего процесса при воздействии события (варианты реакции на воздействие события), связывающие значение некоторого события с некоторым значением или изменением значения признака [16] (например, «при щелчке по камертону ... испытуемый (пациент) должен сообщить, когда он перестанет ощущать вибрацию - в этот момент следует заметить показания шкалы (от 0 до 8); если значение показателя не превышает «5», то это подтверждает диагноз сенсорной невропатии).

При этом каждый из таких вариантов развития может содержать и связи со значениями воздействующих факторов (например, возрастных или гендерных индивидуальных характеристик или внешних факторов).

Искомый запрос дополнительной информации  $Q$  о значении темпорального признака может относиться не только к текущему моменту  $t_{\text{mom}}$  (моменту принятия решения о запросе дополнительной информации), но и к будущему моменту (в тех случаях, когда в тот момент развитие процесса будет иметь наиболее «показательные» внешние проявления); в редких случаях можно надеяться получить значение признака, которое было у объекта ранее (между  $t_0$  и  $t_{\text{mom}}$ ), а момент наблюдения некоторого запрашиваемого события  $t_u$  может предшествовать началу развития процесса.

## 2.2 Алгоритм ранжирования дифференциаторов по времени, влиянию событий и характеристикам системы

Для удобства решения задачи запроса дополнительной информации, направленного на сокращение множества гипотез о диагнозе, предлагается среди запрашиваемых признаков (названия которых обозначим  $f_{\text{ex}}\text{Name}_i$ , а значения -  $f_{\text{ex}}\text{Value}_i$ ) различать *дифференциаторы* множества рассматриваемых гипотез и *антидифференциаторы* (рассматриваемого множества). *Дифференциатор* – тот признак, каждое значение которого сокращает множество гипотез; *антидифференциатор* – такой признак, ни одно из значений которого не сокращает множество гипотез; «частичные» *дифференциаторы* - остальные признаки (которые могут быть ранжированы, например, по доле значений, которые сократят множество гипотез).

Для решения этой задачи с использованием знаний о внутренних процессах и влиянии на них внешних событий предлагаются следующие этапы алгоритма выявления среди известных измеряемых признаков дифференциаторов и «частичных» дифференциаторов.

### 2.2.1 Шаг 1. Предобработка БЗ и построение «рабочей» модели

Из БЗ KB выделить фрагмент KB', содержащий предложения о свойствах ситуаций (диагностируемого объекта) каждого класса (диагноза)  $\Delta_{ij}$  из множества гипотез  $H_{R,KB} = \{\Delta_{i1}, \dots, \Delta_{im}\}$ ; инвертировать фрагмент KB', т.е. построить «рабочую» модель признаков гипотез  $KB'^{\text{Inv}}$  как множество «троек» <признак, «характерные» значения, диагноз> и пар <причина, диагноз>, чтобы в модели рассматривать только *подмножество признаков*  $f_{\text{ex}} \in F_{\text{ex}}$ , факторов и событий  $f_{\text{ev}} \in F_{\text{ev}}$ , которые *связаны с гипотезами* из  $H_{R,KB}$ .

Пример построения «рабочей» модели признаков. Пусть KB' содержит для класса-диагноза «Дифтерийный конъюнктивит» следующие признаки:

выделение из глаз. Присутствие = «имеется»;

выделение из глаз. Локализация = «справа И слева»;

выделение из глаз. Характер начала = «острое»;

выделение из глаз. Характер отделяемого с возможными вариантами динамики: 1. мутное в виде хлопьев; 2. серозно-кровянистое 1-3 суток, затем гнойное.

Варианты причины возникновения процесса «Дифтерийный конъюнктивит» таковы:  
 через 2-6 суток после начала заболевания «дифтерия»;  
 через 2-4 суток после контакта с больным дифтерией.

Тогда  $KB^{Inv}$  будет содержать «тройки»:

<Выделение из глаз. Присутствие, «имеется», Дифтерийный конъюнктивит>;

<Выделение из глаз. Характер начала, «острое», Дифтерийный конъюнктивит>;

<Выделение из глаз. Локализация, «справа И слева»; Дифтерийный конъюнктивит>;

<Выделение из глаз. Характер отделяемого, мутное в виде хлопьев»; Дифтерийный конъюнктивит>;

<Выделение из глаз. Характер отделяемого, {«серозно-кровянистое» (1-3 суток); гнойное}, Дифтерийный конъюнктивит>;

<Выделение из глаз. Периодичность, «постоянно», Дифтерийный конъюнктивит>;

и пары

<«Дифтерийный конъюнктивит» ( $t_{beg}$ ), (начало заболевание «дифтерия» ( $t_u$ ),  $t_{beg}$  через 2-6 суток после  $t_u$ )>;

<«Дифтерийный конъюнктивит» ( $t_{beg}$ ), (контакт с больным дифтерией( $t_u$ ),  $t_{beg}$  через 2-4 суток после  $t_u$ )>.

Для некоторых  $f_{ex}Name_i$  из  $KB'$  могут существовать варианты развития процесса, изменённого воздействием события, связывающие: (признак, диагноз, событие, временной интервал начала такого воздействия, результирующее значение признака) или (признак, исходное значение признака, событие, временной интервал начала такого воздействия, результирующее значение признака). Такие предложения должны быть добавлены в «рабочую» модель признаков  $KB^{Inv}$ .

### 2.2.2 Шаг 2. Исключение лишних признаков из «рабочей» модели признаков

Обозначим  $t_{mom}$  текущий момент - точку на *временной* оси (от начала наблюдений  $t_0$  либо от начала развития внутреннего процесса  $t_{beg}$ ), момент принятия решения о запросе дополнительной информации. В «рабочей» модели  $KB^{Inv}$  удалить (или пометить как «неактивные»):

- каждый признак  $f_{ex}Name_i$ , который является *статическим* и его значение  $f_{ex}Value_i$  уже присутствует в  $R$ ;  
 (подразумевается удаление всех «троек» (признак =  $f_{ex}Name_i$ , «характерные» значения, диагноз);
- те признаки, у которых один из вариантов *динамики значений* определён не далее, чем до текущего момента и все его значения  $f_{ex-j}Value_i(t_1, \dots, t_{mom})$ , указывающие на этот вариант развития, уже присутствуют в  $R$ ;
- те возможные факторы и события  $f_{ev}Name_w (\in F_{ev})$ , информация о существовании которых обычно имеет область возможных значений (ОВЗ) = {да, нет}, ОВЗ уже присутствует в  $R$ .

Пример удаления лишних «троек» из «рабочей» модели признаков.

Пусть  $t_{mom} = 12.04$ ;  $t_0 = 10.04$ ;  $R$  содержит:

выделение из глаз. Присутствие = «имеется»;

выделение из глаз. Локализация = «справа И слева»;

выделение из глаз. Характер начала = «острое»;

выделение из глаз. Характер отделяемого (12.04) = «серозно-кровянистое»;

выделение из глаз. Периодичность = «постоянно»;

контакт с больным дифтерией (8.04), ...

Это позволяет удалить из  $KB^{Inv}$  «тройки»:

<Выделение из глаз. Присутствие, «имеется», Дифтерийный конъюнктивит>;

<Выделение из глаз. Характер начала, «острое», Дифтерийный конъюнктивит>.

Оставшиеся в «рабочей» модели  $KB^{Inv}$  признаки считаются возможными дифференциаторами.

### 2.2.3 Шаг 3. Поиск и исключение антидифференциаторов

Искать в  $KB^{Inv}$  те признаки, у которых:

- «тройки» (признак, «характерные» значения, диагноз) присутствуют для каждого из  $m$  диагнозов  $\{\Delta_{i1}, \dots, \Delta_{im}\} = N_{R,KB}$ , т.е.  $m$  раз, при этом в каждой «тройке» одинаковы «характерные» значения:  
либо одно и то же множество вариантов значений (тот же диапазон значений, то же единственное значение) статического признака;  
либо с каждой гипотезой:  $\Delta_{i1}, \dots, \Delta_{im}$  связаны одинаковые варианты динамики значений признака;  
либо в  $R$  отсутствуют и не могут быть получены значения динамического признака до текущего момента  $t_{mom}$ , но начиная с «текущего момента» (точки на временной оси от начала развития внутреннего процесса), с каждой гипотезой:  $\Delta_{i1}, \dots, \Delta_{im}$  связаны одинаковое значение признака или одинаковый диапазон значений или одинаковые варианты динамики значений во все последующие моменты времени;
- те возможные события  $f_{ex}Name_w$  в момент  $t_v \leq t_0$ , которые присутствуют для каждого из  $m$  диагнозов  $\{\Delta_{i1}, \dots, \Delta_{im}\} = N_{R,KB}$ , считать антидифференциаторами. Их можно удалить из  $KB^{Inv}$  («рабочей» модели признаков).

Примечание. Если все признаки  $KB^{Inv}$  оказались антидифференциаторами, важно сигнализировать о неполноте БЗ и завершить работу (алгоритма). Если НЕ все признаки  $KB^{Inv}$  оказались антидифференциаторами, то в  $KB^{Inv}$  содержатся признаки, которые ещё не имеют значений (отсутствуют в  $R_{ex}(t_0, \dots, t_k)$ ).

### 2.2.4 Шаг 4. Поиск дифференциатора

Искать в  $KB^{Inv}$  признак  $f_{ex}Name_i$ , который для всех гипотез имеет разное значение из допустимого множества значений. Типичные поисковые запросы таковы.

А) Искать дифференциатор - статический признак с «качественными» значениями. Среди статических признаков  $\{f_{ex}Name_i\}$  (в  $KB^{Inv}$ ), имеющих в качестве ОВЗ множество значений, искать  $f_{ex}Name_i$ , который для всех  $m$  гипотез имеет разное значение (из этого допустимого множества значений  $\{f_{ex}Value_{ij}\}, j=1, n$ ).

Б) Искать дифференциатор - статический признак с диапазонами значений. Среди статических признаков  $\{f_{ex}Name_i\}$  (в  $KB^{Inv}$ ), имеющих в качестве ОВЗ один или совокупность числовых диапазонов, найти признак  $f_{ex}Name_i$ , который для всех гипотез имеет непересекающийся поддиапазон значений ( $minValue_j - maxValue_j$ ), т.е.  $minValue_{j+1} > maxValue_j$ ,  $minValue_j > maxValue_{j-1}$ .

В) Искать дифференциатор - динамический признак. Среди признаков  $\{f_{ex}Name_i\}$ , имеющих в любой момент (или период) времени в качестве ОВЗ множество значений, искать  $f_{ex}Name_i$ , который в некоторый момент  $t_v$  (или период  $\langle t_v, t_{v+1} \rangle$ ) времени,  $\geq$  «текущего» момента  $t_{mom}$ , для всех  $n$  гипотез имеет разное значение  $f_{ex}Value_i(t_v)$  из допустимого множества значений  $\{f_{ex}Value_{ij}\}$ .

Среди признаков  $\{f_{ex}Name_i\}$ , имеющих в любой момент (или период) времени в качестве ОВЗ числовые диапазоны, искать признак  $f_{ex}Name_i$ , который в некоторый момент  $t_v$  (или период  $\langle t_v, t_{v+1} \rangle$ ) времени,  $\geq$  «текущ. момента»  $t_{mom}$ , для каждой из  $n$  гипотез имеет такой поддиапазон значений ( $minValue_j - maxValue_j$ ),  $j=1, n$ , что множества таких  $n$  поддиапазонов не пересекаются.



Искать те возможные события  $f_{ev}Name_w$  ( $\in F_{ev}$ ), запрос  $Q$  о существовании (или значении) которых в момент  $t_v \leq t_0$  может иметь ответ (AQ) в «текущий момент»  $t_{mom}$  и этот ответ имеет разное значение для всех  $m$  гипотез-диагнозов.

Далее оценить дифференциатор, если он был найден. Ориентируясь на предложения <Признак, исходное значение признака, Событие, временной интервал начала такого воздействия, результирующее значение признака> в KB<sup>Inv</sup>, понять является ли дифференциатор «подверженным влиянию». Таковым следует считать дифференциатор  $f_{ex}Name_i(t_v)$ , про который известно, что на его значение влияют события  $f_{ev}Name_w$  (например, приём внутрь лекарственных средств или других веществ) и известен временной интервал начала такого воздействия.

Если для каждого такого «подверженного влиянию» дифференциатора  $f_{ex}Name_i$  ответы (AQ) на запросы  $Q$  о таких событиях в моменты  $t_v \leq t_{mom}$  отсутствуют в  $R_{ex}(t_0, \dots, t_k)$ , то  $f_{ex}Name_i$  считать дифференциатором совместно с AQ и при выполнении условий: AQ = нет либо AQ = да, но  $f_{ex}Value_i(t_w)$  в  $t_w \geq t_{mom}$  даже с учётом влияния события имеет для всех гипотез разное значение (или поддиапазон значений) из ОВЗ.

Примечание. В рабочей модели KB<sup>Inv</sup> описание связей признаков с диагнозами может быть сгруппировано по вариантам значений. Найти дифференциатор - значит найти  $f_{ex}Name_i$ , который для всех  $n$  гипотез имеет разное значение из допустимого множества значений  $\{f_{ex}Value_{ij}\}$ .

Если множество дифференциаторов =  $\emptyset$ , применить шаг 5. После шага 4 (Если множество дифференциаторов =  $\emptyset$ ) в KB<sup>Inv</sup> содержатся «частичные» дифференциаторы, которые могут дифференцировать гипотезы, если ответом (AQ) станет «подходящее»/«дифференцирующее» (способное отвергнуть некоторую гипотезу) значение.

### 2.2.5 Шаг 5. Ранжирование частичных дифференциаторов

Используя KB<sup>Inv</sup>, ранжирование признаков  $f_{ex}Name_i$  можно провести по доле значений, которые предположительно сократят множество гипотез (с какой вероятностью в ответе AQ<sub>Σ</sub> окажется значение  $f_{ex}Value_{ij}$ , способное исключить  $k$  из  $n$  гипотез).

Ранжирование статических признаков с качественными значениями (признаки  $\{f_{ex}Name_i\}$ , имеющие ОВЗ множество «качественных» значений) таково: признаки  $\{f_{ex}Name_i\}$  (в KB<sup>Inv</sup>), имеющие ОВЗ множество «качественных» значений, связанные более, чем с одним диагнозом, упорядочить, например, по величине отношения числа разных значений  $f_{ex}Value_{ij}$  признака, связанных только с одним из этих диагнозов, к числу возможных значений. (Далее упорядочить аналогично признаки, у которых разные значения связаны только с двумя диагнозами (если их число в рабочей модели >2) и т.д.).

Ранжирование статических признаков  $\{f_{ex}Name_i\}$ , имеющих ОВЗ множество числовых интервалов, таково: все такие признаки  $\{f_{ex}Name_i\}$ , связанные более, чем с одним диагнозом, упорядочить, например, по величине отношения длины интервала ( $f_{ex,max}Value_{ij} - f_{ex,min}Value_{ij}$ ), связанного только с одним из этих диагнозов, к длине ОВЗ ( $f_{ex,max}Value - f_{ex,min}Value$ ).

Если к шагу 5 в KB<sup>Inv</sup> множество признаков  $\neq \emptyset$ , то хотя бы один частичный дифференциатор будет предложен (признак, значение которого, возможно, сократило множество гипотез).

Примечание. При наличии в рабочей модели KB<sup>Inv</sup> признаков с динамическими значениями признаки в KB<sup>Inv</sup> могут быть сгруппированы по их вариантам динамики; кроме того, в рабочей модели KB<sup>Inv</sup> значения признаков могут быть сгруппированы по периодам динамики (которым сопоставлены диагнозы, при которых такой вариант возможен); кроме того,

рабочая модель KB<sup>Inv</sup> может быть организована по значениям признаков в разные подынтервалы периодов динамики.

## 2.3 Применение алгоритма

Представленный алгоритм выявления дифференциаторов среди известных признаков (в соответствии с постановкой задачи запроса) применяется многократно в процессе дифференциации. При наличии  $n$  гипотез о диагнозе обращение к алгоритму даст ответ на вопрос: «Какой признак запросить, чтобы ответ на него помог отвергнуть хотя бы одну гипотезу?». Если после работы алгоритма запрос признака привел к получению ответа (значение указанного признака в указанный момент времени), то этим ответом в соответствии с рекомендацией алгоритма может быть отвергнута указанная гипотеза либо, поскольку представление о ситуации уточняется, запускается процесс генерации нового множества гипотез, ожидаемо меньшего. В результате оказывается  $n' = n-1$  гипотез (или даже меньше). Если  $n' > 1$ , то необходимо очередное обращение к этому алгоритму. В идеале цикл обращений завершается, когда  $n' = 1$ .

## Заключение

При решении задач дифференциальной диагностики и некоторых других подклассов задачи распознавания [11] специалисты вынуждены решать задачу определения того, какой информации не хватает, чтобы уточнить множество решений или сократить его до единственного.

Предложенный алгоритм выявления дифференциаторов даёт возможность объективного определения того, какую информацию (с перечисленными особенностями) следует запросить, чтобы уточнить множество решений. Такой алгоритм применим для решения задачи запроса дополнительной информации при распознавании ситуаций в различных областях деятельности. Наличие такого алгоритма крайне важно при проектировании систем поддержки диагностических решений в ситуациях, не имеющих в начальный момент всей необходимой информации. При этом можно ожидать, что наличие в ПрО дополнительных знаний о зависимостях между внешними признаками развития процессов и нарушениями или развитием с отклонениями от нормы (уровни доказательности и убедительности) позволит упростить некоторые шаги этого алгоритма для конкретных ПрО.

## Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ (проекты №№ 15-07-03193, 17-07-00956).

## Список источников

- [1] **Ле, Н.В.** Модель представления знаний при создании медицинской экспертной системы дифференциальной диагностики / Н.В. Ле, В.А. Камаев, Д.П. Панченко, О.А. Трушкина // Известия Волгоградского государственного технического университета. - 2014. - № 6. - С. 42-50.
- [2] **Власов, А.И.,** Архитектура автоматизированной экспертной системы диагностики технического состояния турбогенераторов / А.И. Власов, Э.В. Манукянц, Э.В. Мысловский // Инженерный вестник: электронный научно-технический журнал. - 2012. - №9. – [http://iu4.ru/publ/2012\\_ing\\_vest\\_09\\_04.pdf](http://iu4.ru/publ/2012_ing_vest_09_04.pdf).
- [3] **Литвиненко, В.И.** Некоторые аспекты категорного представления информации в лечебно-диагностических системах / В.И. Литвиненко, Н.А. Соколова, А.П. Бень // Вестник ХГТУ. - 1997. - № 1. - С. 101 -104.

- [4] **Lemaire, J.B.** Effectiveness of the Quick Medical Reference as a diagnostic tool / J.B. Lemaire, J.P. Schaefer, L.A. Martin, et al. // CMAJ - 1999. - Vol. 161 (6). - P. 725-728.
  - [5] **Miller, R.A.** INTERNIST-1: An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine / R.A. Miller, H. Pople, J. Myers // N Engl J Med. - 1982. - Vol. 307. - P. 468-476.
  - [6] **Patil, R.S.** Modelling Knowledge of the Patient in Acid-base and Electrolyte Disorders / R.S. Patil, P. Szolovits, W.B. Schwartz // in Szolovits, P. (Ed.). Artificial Intelligence in Medicine, - Boulder, CO: Westview Press, 1982. - P. 191-226.
  - [7] **Pople, H.E.** Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics / **H.E. Pople, Jr.** // in Szolovits, P. (Ed.) Artificial Intelligence in Medicine, - Boulder, CO: Westview Press, 1982. P. 119-185.
  - [8] **Soltan, R.A.** Diagnosis of Some Diseases in Medicine via computerized Experts System / R.A. Soltan, M.Z. Rashad, B. El-Desouky // International Journal of Computer Science & Information Technology - 2013. - Vol. 5(5). - P. 79-90.
  - [9] **Тельнов, Ю.Ф.** Интеллектуальные информационные системы / Ю.Ф. Тельнов. - М.: МЭСИ, 2004. - 246 с.
  - [10] **Davis, R.** Production Rules as a Representation for a Knowledge Based Consultation Program / R. Davis, B.G. Buchanan, E.H. Shortliffe // Chapter 5 in Clancey, W.J. and Shortliffe E.H. (Ed.) Readings in Medical Artificial Intelligence: The First Decade, - Addison Wesley, Reading, MA, - 1984. - P. 98-130. - <http://people.dbmi.columbia.edu/~ehs7001/Clancey-Shortliffe-1984/Ch5.pdf>.
  - [11] **Клещёв, А.С.** Онтология задач интеллектуальной деятельности / А.С. Клещёв, Е.А. Шалфеева // Онтология проектирования. - 2015. - Т. 5. №2(16). - С. 179-205. DOI:10.18287/2223-9537-2015-5-2-179-205.
  - [12] **Denekamp, Y.** TiMeDDx — A multi-phase anchor-based diagnostic decision-support model / Y. Denekamp, M. Peleg // Journal of Biomedical Informatics. - 2010. - Vol. 43. - P. 111-124.
  - [13] **Кобринский, Б.А.** Консультативные интеллектуальные медицинские системы: классификации, принципы построения, эффективность / Б.А. Кобринский // Врач и информационные технологии. - 2008. - № 2. - С. 38-47.
  - [14] **Черняховская, М.Ю.,** Формальное представление знаний о конъюнктивитах (издание второе, исправленное и дополненное) / М.Ю. Черняховская, Ф.М. Москаленко, В.Я. Мельников, В.И. Негода, Л.П. Догадова. - Владивосток: ИАПУ ДВО РАН, 2009. - 56 с.
  - [15] **Inzartsev, A.** Application of Artificial Intelligence Techniques for Fault Diagnostics of Autonomous Underwater Vehicles / A. Inzartsev, A. Pavin, A. Kleschev, V. Gribova, G. Eliseenko // Proc. of the OCEANS 2016 MTS/IEEE Conference & Exhibition, (September 19-23, 2016, Monterey, California, USA), ISBN DVD: 978-1-5090-1538-2.
  - [16] **Клещёв, А.С.** Модель онтологии предметной области «Медицинская диагностика». Часть 1. Неформальное описание и определение базовых терминов / А.С. Клещёв, М.Ю. Черняховская, Ф.М. Москаленко // Журнал НТИ - Серия 2. - 2005. - № 12. - С. 1-7.
  - [17] **Боргест, Н.М.** Ключевые термины онтологии проектирования: обзор, анализ, обобщения / Н.М. Боргест // Онтология проектирования. - 2013. - № 3 (9). С. 9-31.
-

## THE METHOD OF SOLUTION OF THE PROBLEM OF ADDITIONAL INFORMATION REQUEST

A.S. Kleschev, V.V. Gribova, E.A. Shalfeeva

*Institute for Automation Control Processes of the FEB RAS, Vladivostok, Russia*  
kleschev@dvo.ru, gribova@iacp.dvo.ru, shalf@iacp.dvo.ru

### Abstract

The article considers a problem of additional information request, usually being solved in diagnostics process for reduction of number of diagnosis hypotheses. In practice during solution of these two tasks it is important to consider all types of the cause-and-effect relations, variability of signs values for a time, influence of events on dynamics of values of signs. The offered algorithm of solution of this task is based on choosing differentiators among the known measurable signs. Identification of differentiators is made with use of knowledge of internal processes developing and external events' influence on them. We will apply such algorithm to the solution of a problem of request of additional information for situations recognition, classification and diagnostics in various spheres of activity.

**Key words:** domain ontology, diagnostics problem, set of hypotheses, additional information request problem, differentiating sign, task solution algorithm.

**Citation:** Kleshchev AS., Gribova VV., Shalfeeva EA. The method of solution of the problem of additional information request. *Ontology of designing*. 2017; 7(3): 310-322. DOI: 10.18287/2223-9537-2017-7-3-310-322.

### References

- [1] **Le NV., Kamayev VA., Panchenko DP., Trushkina OA.** Model of representation of knowledge during creation of medical expert system of differential diagnostics. [In Russian]. *News of the Volgograd state technical university*, 2014; 6: 42-50.
- [2] **Vlasov AI., Manukyants EV., Myslovsky EV.** Architecture of the automated expert system of diagnostics of technical condition Turbogenerators [In Russian]. *Engineering messenger: online scientific and technical magazine*, 2012; 9: [http://iu4.ru/publ/2012\\_ing\\_vest\\_09\\_04.pdf](http://iu4.ru/publ/2012_ing_vest_09_04.pdf).
- [3] **Litvinenko VI., Sokolova HA., Ben AP.** Some aspects of category submission of information in medical and diagnostic systems [In Russian]. *HGTU bulletin*, 1997; 1: 101-104.
- [4] **Lemaire JB., Schaefer JP., Martin LA., et al.** Effectiveness of the Quick Medical Reference as a diagnostic tool *CMAJ*; 1999. 161 (6): 725-728.
- [5] **Miller RA., Pople H., and Myers J.** INTERNIST-1: An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine. *N Engl J Med.*; 1982. 307: 468-476.
- [6] **Patil RS., Szolovits P., Schwartz WB.** Modelling Knowledge of the Patient in Aced-base and Electrolyte Disorders. In Szolovits, P. (Ed.) *Artificial Intelligence in Medicine*, - Boulder, CO: Westview Press; 1982: 191-226.
- [7] **Pople HE., Jr.** Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics. In Szolovits, P. (Ed.) *Artificial Intelligence in Medicine*, - Boulder, CO: Westview Press; 1982: 119-185.
- [8] **Soltan RA., Rashad MZ., El-Desouky B.** Diagnosis of Some Diseases in Medicine via computerized Experts System. *International Journal of Computer Science & Information Technology*; 2013. 5(5): 79-90.
- [9] **Telnov YF.** Intellectual information systems. [In Russian]. Moscow: MESI, 2004.
- [10] **Davis R., Buchanan BG., Shortliffe EH.** Production Rules as a Representation for a Knowledge Based Consultation Program. Chapter 5 in Clancey W.J. and Shortliffe E.H. (Ed.) *Readings in Medical Artificial Intelligence: The First Decade*. Addison Wesley, Reading, MA, 1984: 98-130: <http://people.dbmi.columbia.edu/~ehs7001/Clancey-Shortliffe-1984/Ch5.pdf>.
- [11] **Kleshchev AS., Shalfeeva EA.** Ontology of problems of intellectual activity [In Russian]. *Ontology of Designing*, 2015; 5(2): 179-205. DOI:10.18287/2223-9537-2015-5-2-179-205.
- [12] **Denekamp Y., Peleg M.** TiMeDDx — A multi-phase anchor-based diagnostic decision-support model. *Journal of Biomedical Informatics*, 2010; 43: 111-124.

- [13] **Kobrinsky BA.** Advisory intellectual medical systems: classifications, principles of construction. [In Russian]. Doctor and information technologies, 2008; 2: 38-47.
- [14] **Chernyakhovskaya MYu., Moskalenko FM., Melnikov VYa., Negoda VI., Dogadova LP.** Formal representation of knowledge of conjunctivitis (the edition second corrected and added) [In Russian]. Vladivostok: IAPU FEB RAS, 2009.
- [15] **Inzartsev A., Pavin A., Kleshev A., Gribova V., Eliseenko G.** Application of Artificial Intelligence Techniques for Fault Diagnostics of Autonomous Underwater Vehicles // Proceedings of the OCEANS 2016 MTS/IEEE Conference & Exhibition (Monterey, California, USA, September 19-23, 2016), ISBN DVD: 978-1-5090-1538-2.
- [16] **Kleshchev AS., Chernyakhovskaya MYu., Moskalenko FM.** Model of ontology of subject domain "Medical diagnostics". Part 1. Informal description and definition of basic terms. [In Russian]. NTI. A Series 2, 2005; 12: 1-7.
- [17] **Borgest NM.** Keywords of ontology of designing: review, analysis, generalization [In Russian]. *Ontology of Designing*, 2013, № 3(9): 9-31.

## Сведения об авторах



**Клещев Александр Сергеевич**, 1940 г. рождения. Окончил математико-механический факультет Ленинградского государственного университета в 1964 г., д.ф.-м.н. (1990). Главный научный сотрудник лаборатории интеллектуальных систем Института автоматизации и процессов управления Дальневосточного отделения РАН, профессор, заслуженный деятель науки РФ. В списке научных трудов более 300 работ в области искусственного интеллекта, информатики, медицинской и биологической кибернетики.

**Alexander Sergeevich Kleshev** (b. 1940) graduated from the Leningrad State University in 1964, Professor's degree (1990). He is Chief Researcher at lab. of intellectual systems in the Institute for Automation & Control Processes of the FEB RAS. He is co-author of more than 300 publications in the fields of biological and medical cybernetics, informatics and AI.



**Грибова Валерия Викторовна**, 1965 г. рождения. Окончила Ленинградский политехнический институт по специальности "прикладная математика" (1989), д.т.н. (2007). Заведующая лабораторией интеллектуальных систем Института автоматизации и процессов управления Дальневосточного отделения РАН, заместитель директора по научной работе, вице-президент Российской ассоциации искусственного интеллекта. В списке научных трудов более 200 работ в области искусственного интеллекта, проблемно-ориентированных систем, основанных на знаниях, специализированных программных моделей и систем.

**Valeriya Victorovna Gribova** (b.1965) graduated from the Leningrad Polytechnic University in 1989, Professor's degree (2007). She is a Head of lab. of intellectual systems in the Institute for

Automation & Control Processes of the FEB RAS, Research Deputy Director, an expert of Analytic Center in Government of Russian Federation, a Vice-President of Russian Association of Artificial Intelligence. She is co-author of more than 200 publications in the fields of AI, informatics, Program models, technologies and systems.



**Шалфеева Елена Арефьевна**, 1967 г. рождения. Окончила математический факультет Дальневосточного государственного университета по специальности "прикладная математика" (1989), к.т.н. (2000). Старший научный сотрудник лаборатории интеллектуальных систем Института автоматизации и процессов управления Дальневосточного отделения РАН, доцент по специальности. В списке научных трудов более 80 работ.

**Elena Arefyevna Shalfeeva** (b.1967) graduated from the Far Eastern State University (Vladivostok) in 1989, PhD (2000). She is Senior Researcher at lab. of intellectual systems in the Institute for Automation & Control Processes of the FEB RAS, lecture. She is co-author of more than 80 publications in the fields of Program models and systems and AI.