

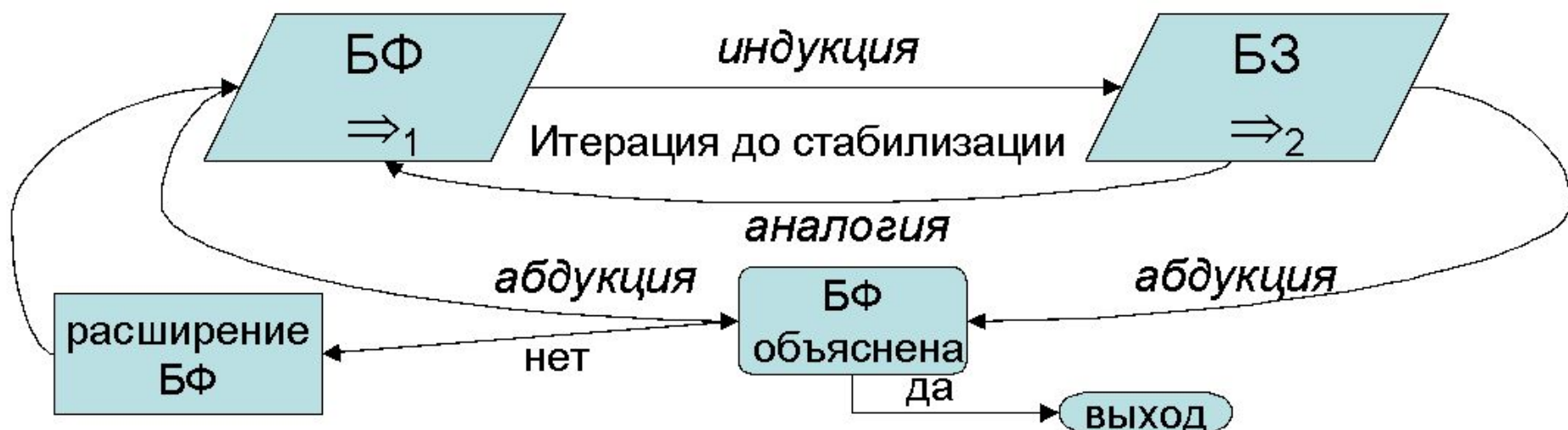
Интеллектуальная система типа ДСМ для анализа клинических данных

**Панкратова Е.С., Добрынин Д.А.
(ВИНИТИ РАН и РГГУ),
Михайлова И.Н.(РОНЦ
им. Н.Н.Блохина РАМН)**

Условия применимости ДСМ-метода

- Возможность структурирования данных и формального определения сходства фактов(из БФ).
- Наличие положительных и отрицательных примеров в БФ.
- Наличие в БФ неявно заданных зависимостей причинно-следственного типа (причины изучаемых эффектов).

Общая схема ДСМ-рассуждений



- (1). *Индукция*. Порождение гипотез о причинах наличия (отсутствия) свойств объектов ((\pm)-причинах) (\Rightarrow_2 , БЗ) на основе анализа сходства примеров из БФ (\Rightarrow_1). ППВ-I.
- (2). *Аналогия*. Порождение гипотез о (\pm)-эффектах (\Rightarrow_1 , БФ) на основе структурной аналогии с использованием гипотез о причинах (\Rightarrow_2 , БЗ), полученных на предыдущем этапе. ППВ-II.
- (3). Итерация (1)&(2) до стабилизации.
- (4). *Абдукция*. Объяснение исходной БФ (\Rightarrow_1) на основе полученных гипотез о причинах (\Rightarrow_2 , БЗ). АКП.

Особенности ДСМ-метода

1. ДСМ-метод - нестатистический метод анализа данных.
2. Способность порождать полезные гипотезы на малых массивах данных.
3. Возможность учитывать индивидуальные особенности изучаемых объектов исследования.
4. Работает с открытыми массивами данных, указывая на необходимость расширения базы фактов, если таковая возникает, (это осуществляется посредством абдуктивного объяснения базы фактов).

Настройка на предметную область

1. Разработка языка представления данных.
2. Определение понятия «объект» и «свойство» в терминологии ДСМ-метода.
3. Определение аксиом предметной области.
4. Задание операции сходства.
5. Задание отношения вложения.

ИС типа ДСМ для медицины

1. Прогнозирование высокопатогенных типов вируса папилломы человека (ВПЧ) по цитологическим результатам исследования мазков (Кафедра клинической и лабораторной диагностики Российской Медицинской Академии последипломного образования)
2. Диагностика двух заболеваний глаз: дегенеративного ретиношизиса и наследственных витреоретинальных дистрофий (Лаборатория клинической физиологии зрения МНИИ глазных болезней им. Гельмгольца)
3. Диагностика и определение способа лечения «Красной волчанки» по клиническим и лабораторным признакам больного (Отделение нефрологии Городской клинической больницы им. Боткина)
4. Прогнозирование продолжительности жизни больных меланомой и оценка прогностического биохимического маркера S100 (РОНЦ им.Н.Н.Блохина РАМН)

Эффективность

подобных исследований зависит от максимально полного использования совокупного анализа результатов различных обследований больного.

Разработка языка представления данных

Типы данных

1. В кортеже длины «n» (где n- количество элементов списка признаков) указываются **присутствующие качественные признаки**, вместо отсутствующих признаков ставится Λ (пустой элемент):
$$E_i^{(1)} = \langle At_1, \dots, At_i, \dots, At_n \rangle.$$

Пример из задачи №4

$EI^{(1)}$ = Пигментные образования кожи (возможно несколько признаков, отмечаются только присутствующие признаки):

At_1 = 5 и более родинок размерами от 6 мм в диаметре «+»

At_2 = 50 и более родинок размерами от 3 мм в диаметре

At_3 = Наличие одного или более атипичного (диспластического) невуса «+»

At_4 = Врожденные невусы

At_5 = Гигантский врожденный невус

At_6 = Меланома (в анамнезе)

$EI^{(1)} = \langle At_1, \Lambda, At_3, \Lambda, \Lambda, \Lambda \rangle$

ВОЗМОЖНЫХ качественных признаков:

$E1^{(2)} = At_i, i \in \{1, \dots, n\}$, n -число признаков в списке.

Пример: $E1^{(2)}$ = тип роста первичной опухоли

At_1 = поверхностно распространяющаяся меланома,

At_2 = узловая меланома, «+»

At_3 = меланома по типу злокачественного лентиго,

At_4 = акролентигинозная меланома,

выбирается один из атрибутов, например,

$E1^{(2)} = At_2$ = узловая меланома

3. В n -элементном кортеже указываются **все присутствующие (+) и все отсутствующие (-) признаки** из заранее разработанного списка, наличие или отсутствие остальных признаков считается неизвестным (τ).

$$E1^{(3)} \langle \langle At_1, k_1 \rangle, \dots, \langle At_i, k_i \rangle, \dots, \langle At_n, k_n \rangle \rangle,$$

$k_i \in \{+, -, \tau\}$, где при $k_i = \tau$, $\langle At_i, k_i \rangle = \Lambda$

2. Пример из задачи (2) :

$E1^{(3)}$ =стекловидное тело:

At_1 =«плавающие помутнения +»,

At_2 = «нитчатые помутнения – »,

At_3 = «мембраны τ »,

At_4 =выраженная зернистая деструкция +»,

At_5 =преритинальные пленки +»,

At_6 = «шварты τ »,

At_7 = «задняя отслойка стекловидного тела +»,

$E1^{(3)}$ = $\langle\langle At_1,+ \rangle, \langle At_2,- \rangle, \Lambda, \langle At_4,+ \rangle, \langle At_5,+ \rangle, \Lambda, \langle At_7,+ \rangle\rangle$

4. Указывается конкретный признак иерархической структуры со знаком «+» или «-», («t» по умолчанию)

Пример из задачи (2):

изменение сосудов сетчатки (да – нет, t)

At_1 – изменение артерий

At_1^1 – сужение артерий

$At_{1.1}^{1.1}$ – расширение артерий «-»

$At_{1.2}^{1.2}$ – извитость артерий

$At_{1.3}^{1.3}$ – изменение вен «+»

At_2 – сужение вен

$At_{2.1}^{2.1}$ – расширение вен «+»

$At_{2.2}^{2.2}$ – извитость вен

$At_{2.3}^{2.3}$ – новообразованные сосуды

At_3 – муфты по ходу сосудов

At_4 – шварта сосудов

$EI^{(5)} = \langle At_{11.1}, At_{11.1.1}, At_{11.1.2}, \dots \rangle$

5. Норма признака или интервал отклонения признака от нормы с указанием направления отклонения

$EI^{(4)} = \langle At, dir, k \rangle$, где $dir \in \{\downarrow, N, \uparrow\}$, N – норма признака, \downarrow – отклонение от нормы в сторону уменьшения, \uparrow – отклонение от нормы в сторону увеличения, а $k \in \{1, 2, 3, 4, 5\}$

В задаче (2): $EI = \langle A\text{-волна ЭРГ}, \downarrow, 2 \rangle$ или $EI^{(4)} = \langle A\text{-волна ЭРГ}, N, \rangle$.

ИС типа ДСМ для анализа клинических данных больных меланомой

Объект, соответствующий истории болезни конкретного больного, представляет собой кортеж из 62-х элементов.

$$O_i = \langle EI_1, \dots, EI_{62} \rangle$$

Каждый элемент этого кортежа соответствует конкретному признаку. В соответствии с языком представления данных каждый элемент объекта принадлежит одному из 3-х типов.

Свойством является прогноз продолжительности жизни.

Определение аксиом предметной области

Все перечисленные признаки, входящие в объект, возможно разделить с учетом знаний о предметной области на три группы:

- 1) Группа необходимых признаков (конъюнктивных) – признаки, без наличия которых гипотеза не имеет смысла, например, нельзя прогнозировать результат лечения без наличия признака **терапия**.
- 2) Группа существенных признаков (дизъюнктивных) – признаки без наличия хотя бы одного из них гипотеза не имеет смысла.
- 3) Группа сопутствующих признаков - признаки не входящие в группы 1) и 2), например: пол, возраст, артериальное давление и т.д.

Аксиома

Результат операции сходства считается гипотезой, если в нее входит хотя бы один из признаков группы 2) и все признаки группы (1).



Конъюнктивные фильтры

Задача №3 – способ лечения

Задача №4 – S100



Настройка системы на эксперимент

1. Выбор стратегии: простой метод сходства, метод сходства с запретом на контрпримеры отдельно для (+) и (-) примеров.
2. Подбор нужного количества родителей.
3. Настройка фильтра конъюнктивных признаков.
4. Настройка фильтра дизъюнктивных признаков.

Критерий оценки подбора параметров и стратегии эксперимента

Процедура «доопределение по одному»: последовательно каждому объекту выборки присваивается значение «т», производится доопределение этого объекта средствами ДСМ-системы с выбранными параметрами и сравнивается доопределенное значение с существующим.

ИС типа ДСМ для анализа клинических данных больных меланомой

Исходные данные:

38 больных, из них у 12 продолжительность жизни больше 5 лет.

Результаты эксперимента:

У 21-го больного продолжительность жизни определена верно, у 3-х ошибочно, остальные недоопределены.

Результаты эксперимента

Все порожденные положительные гипотезы-причины (продолжительности жизни больше 5 лет), помимо прочих атрибутов, включают значение S100 меньше 0,120 нг/мл (лабораторная норма РОНЦ им. Н.Н.Блохина), а все отрицательные (продолжительность жизни меньше 5 лет) - значение S100 больше 0,120 нг/мл.

Интеллектуальная система типа ДСМ является инструментом поддержки медицинских исследований со сложноструктурированными данными и множеством фактов, необозримых без использования компьютерных технологий.

ИС ДСМ относятся к evidence based medicine

Распространенный сейчас термин evidence based medicine , к сожалению, переведен как «доказательная медицина», а не как «медицина, основанная на очевидных фактах».

Методы, относящиеся к evidence based medicine, основаны на установленных фактах и используют эти факты как аргументы при принятии решений в медицине, в частности, для диагностики по клиническим данным.

Методы, относящиеся к evidence based medicine, основаны на установленных фактах и используют эти факты как аргументы при принятии решений в медицине, в частности, для диагностики по клиническим данным.

Смысл этого термина состоит в том, что решения, принимаемые врачами, должны быть аргументированными (но не доказанными – ведь авторы говорят о вероятных прогнозах, т.е. о гипотезах!), что означает, что принимаемые решения используют аргументы, извлеченные из клинических данных.

Параллельные версии программ были созданы студентками 5-го курса О.Шестерниковой и А. Волковой.