Задача об изоморфизме множеств эмпирических зависимостей, порождаемых из двух обучающих выборок прецедентов.

Типы данных (описаний прецедентов):

- структурные (*Str*)
  - множества
  - графы
  - цепочки символов конечного алфавита
- числовые (*Num*)
  - векторы числовых значений (существенных параметров)
- кортежи вида <*Str*,*Num*>

Операция сходства: бинарная алгебраическая операция $\otimes$, для каждого выбранного типа представления данных - описаний прецедентов исходной обучающей выборки Ω в виде множеств, графов, цепочек символов конечного алфавита, … - определяемая, соответственно, для множеств – как их пересечение, для графов - как выделение множества всех максимальных общих подграфов двух (одноэлементных множеств) графов, для цепочек символов конечного алфавита – как выделение множества всех максимальных общих подцепочек двух (одноэлементных множеств) цепочек и т.п. , удовлетворяющая условиям

- $a \otimes a = a$
- $a \otimes b = b \otimes a$
- $(a \otimes b) \otimes c = a \otimes (b \otimes c) = a \otimes b \otimes c$

Оператор *Clos* $\quad$ $\boldsymbol{Clos}^{\otimes}: 2^{\Omega} \rightarrow 2^{\Omega}$

по заданному сходству V нескольких (формирующих некоторое непустое подмножество $\Omega' \subseteq \Omega$) прецедентов исходной обучающий выборки Ω выделяет все те ее (оставшиеся – т.е. не включенные в Ω') элементы (формирующие множество $\Omega'' \subseteq \Omega$), такие, что сходство всех прецедентов из множества $\Omega' \cup \Omega''$ совпадает с заданным V (уже полученным из «стартового» подмножества прецедентов Ω').

Задача 1: показать, что ***Clos*** есть оператор замыкания на множестве Ω , т.е., что для любых подмножеств прецедентов $\boldsymbol{O}$, $\boldsymbol{O_i}$ и $\boldsymbol{O_j}$ из 2^{Ω} выполняются следующие три условия

(i) $\boldsymbol{O} \subseteq \boldsymbol{Clos}^{\otimes}(\boldsymbol{O})$
(ii) $\boldsymbol{O_i} \subseteq \boldsymbol{O_j}$ влечет $\boldsymbol{Clos}^{\otimes}(\boldsymbol{O_i}) \subseteq \boldsymbol{Clos}^{\otimes}(\boldsymbol{O_j})$
(iii) $\boldsymbol{Clos}^{\otimes}(\boldsymbol{Clos}^{\otimes}(\boldsymbol{O})) = \boldsymbol{Clos}^{\otimes}(\boldsymbol{O})$.

Диаграмма эмпирических зависимостей: по заданной исходной выборке Ω оператор ***Clos*** порождает множество специальных ее подмножеств, каждое из которых содержит *те и только те* прецеденты из Ω, сходство которых совпадает с заданным. На множестве всех таких подмножеств можно задать отношение частичного порядка (как вложимость подмножеств прецедентов исходной выборки Ω). Диаграмму порождаемого таким образом частичного порядка $\subseteq$ будем обозначать как $\mathbf{D}(\Omega)$.

Задача 2: Является ли $\mathbf{D}(\Omega)$ решеткой?

Задача 3: Можно ли достроить $\mathbf{D}(\Omega)$ до решетки $\mathbf{D}^{\Delta}(\Omega)$? Как именно?

Задача 4: Является ли решетка $\mathbf{D}^{\Delta}(\Omega)$ дистрибутивной?

Ёмкость (число вершин) решетки $\mathbf{D}^{\Delta}(\Omega)$: в исходной обучающей выборке Ω имеется n элементов. Оценить размеры множества числа вершин диаграммы $\mathbf{D}(\Omega)$.

Задача 5: Сколько вершин в диаграмме $\mathbf{D}(\Omega)$ и в решетке $\mathbf{D}^{\Delta}(\Omega)$?

Проблема изоморфизма решеток: имеем две выборки прецедентов Ω_1 и Ω_2 . Изоморфны ли порождаемые ими решетки $\mathbf{D}^{\Delta}(\Omega_1)$ и $\mathbf{D}^{\Delta}(\Omega_2)$?

Задача 6: Изоморфны ли диаграммы $\mathbf{D}(\Omega_1)$ и $\mathbf{D}(\Omega_2)$?

Задача 7: Изоморфны ли решетки $\mathbf{D}^{\Delta}(\Omega_1)$ и $\mathbf{D}^{\Delta}(\Omega_2)$?

Задача 8: Сформулировать необходимые и достаточные условия изоморфизма решеток $\mathbf{D}^{\Delta}(\Omega_1)$ и $\mathbf{D}^{\Delta}(\Omega_2)$.

Задача 9: Сформулировать (полиномиально-) быстро проверяемые достаточные условия изоморфизма (и неизоморфизма) решеток $\mathbf{D}^{\Delta}(\Omega_1)$ и $\mathbf{D}^{\Delta}(\Omega_2)$.

***

Уточнения:

1. Каждый из прецедентов *Exm*, формирующих исходную обучающую выборку - множество ***Exm***, представлен в виде пары

   $$Exm = \langle \boldsymbol{O}, \boldsymbol{P} \rangle ,$$

   где $\boldsymbol{O}$ - есть описание объекта (множество признаков, граф, цепочка символов конечного алфавита, …), а $\boldsymbol{P}$ - непустое подмножество некоторого множества свойств $\boldsymbol{P}_{Exm} = \{P_1, P_2, \ldots P_r\}$.

   Требуется: найти решение ***Задач 1-9***.

2. Пусть $\boldsymbol{P}_{Exm}$ есть одноэлементное множество свойств $\boldsymbol{P}_{Exm} = \{P_1\}$.

   Требуется: найти решение ***Задач 1-9***.

3. Пусть исходная обучающая выборка ***Exm*** есть объединение двух непересекающихся множеств прецедентов $\boldsymbol{Exm}^+$ и $\boldsymbol{Exm}^-$:

   $$\boldsymbol{Exm} = \boldsymbol{Exm}^+ \cup \boldsymbol{Exm}^- ,$$

   первое из которых (будем называть его множеством *примеров*) содержит лишь пары $\langle \boldsymbol{O}^+_i, \boldsymbol{P}^+_i \rangle$, где известно[1], что объект $\boldsymbol{O}^+_i$ обладает множеством свойств $\boldsymbol{P}^+_i$:

   $$\boldsymbol{Exm}^+ = \{\langle \boldsymbol{O}^+_1, \boldsymbol{P}^+_1 \rangle, \langle \boldsymbol{O}^+_2, \boldsymbol{P}^+_2 \rangle, \ldots \langle \boldsymbol{O}^+_{s+}, \boldsymbol{P}^+_{s+} \rangle, \} ,$$

   а второе (множество *контрпримеров*) содержит лишь пары $\langle \boldsymbol{O}^-_i, \boldsymbol{P}^-_i \rangle$, где известно[2], что объект $\boldsymbol{O}^-_i$ обладает множеством свойств $\boldsymbol{P}^-_i$:

[1] Как эмпирический факт.

$$\boldsymbol{Exm}^- = \{<\boldsymbol{O}^-_1,\boldsymbol{P}^-_1>, <\boldsymbol{O}^-_2,\boldsymbol{P}^-_2>, \dots <\boldsymbol{O}^-_{s-},\boldsymbol{P}^-_{s-}>, \}.$$

Сходство примеров Exm_1 и Exm_2 определим как пару

$$<\boldsymbol{O}^+_1 \otimes \boldsymbol{O}^+_2, \boldsymbol{P}^+_1 \cap \boldsymbol{P}^+_2>.$$

Сходство *примеров* Exm_1 и Exm_2 непусто, если оба его компонента непусты. (Аналогичным образом определяется и сходство *контрпримеров*). В диаграмме $\mathbf{D}^+(\boldsymbol{Exm}^+)$ оставим лишь те элементы диаграммы $\mathbf{D}(\boldsymbol{Exm}^+)$, которые не встречаются в диаграмме $\mathbf{D}(\boldsymbol{Exm}^-)$. Аналогичным образом (т.е. сравнением элементов диаграммы $\mathbf{D}(\boldsymbol{Exm}^-)$ с элементами диаграммы $\mathbf{D}(\boldsymbol{Exm}^+)$) сформируем диаграмму $\mathbf{D}^-(\boldsymbol{Exm}^-)$.

Требуется: найти решение ***Задач 1-9***.

4. Пусть (как и выше, в п.3) исходная обучающая выборка ***Exm*** есть объединение двух непересекающихся множеств прецедентов $\boldsymbol{Exm}^+$ и $\boldsymbol{Exm}^-$:

$$\boldsymbol{Exm} = \boldsymbol{Exm}^+ \cup \boldsymbol{Exm}^-.$$

Будем учитывать в каждой из диаграмм $\mathbf{D}^+(\boldsymbol{Exm}^+)$ и $\mathbf{D}^-(\boldsymbol{Exm}^-)$ лишь те сходства, которые не вкладываются ни в один из прецедентов противоположного знака: если пара $<v^{\alpha}, w^{\alpha}>$ (где $\alpha \in \{+,-\}$) есть непустое сходство нескольких примеров из $\boldsymbol{Exm}^{\alpha}$, то неверно, что в $\boldsymbol{Exm}^{-\alpha}$ найдется прецедент $<\boldsymbol{O}^{-\alpha}_0,\boldsymbol{P}^{-\alpha}_0>$, такой, что

$$\boldsymbol{O}^{-\alpha}_0 \otimes v^{\alpha} = v^{\alpha}$$

и

$$w^{\alpha} \subseteq \boldsymbol{P}^{-\alpha}_0.$$

(Это дополнительное условие будем называть *Запретом на КонтрПримеры* - ***ЗКП***).

Требуется: найти решение ***Задач 1-9***.

5. *Дополнительно*.

Операция сходства может быть определена и на векторах числовых значений существенных (с точки зрения анализа соответствующих свойств) параметров. Далее вся схема п.п.1-4 может быть воспроизведена и для этого случая.

Возможная программа действий:

- Теоретико-множественный случай (объекты и свойства – множества «образующих») – разбирается как пример в рамках (1-семестровых) занятий на специальном семинаре.
- Остальные варианты представления данных (графы, цепочки, числовые векторы, кортежи) – варианты для проработки в рамках НИР магистрантов.

---

[2] Как эмпирический факт.