

ОБУЧЕНИЕ В ПРОСТРАНСТВОВОТО НА ВЕРСИИТЕ

Лекция 10

Въведение.

- Пространството на версиите се състои от множеството обобщени и от множеството специализирани модели, които могат да се изградят въз основа на наличните положителни и отрицателни примери.
- Ще се запознаем с един елегантен подход (предложен от Tom M. Mitchell) за изграждане на пространството на версиите чрез използване на положителни и отрицателни примери.
- Обучението в пространството на версиите става бързо при наличието на поток от положителни и отрицателни примери и фиксиран брой от характеристики.
- Използването на *пространството на версиите* дава възможност да се съхранява **нишката** за цялата полезна информация, получена от последователността от обучаващи примери, без необходимост да се помнят самите примери
- Като илюстрация на подхода ще дискутираме възможността да се извличат знания от база с данни.
- Разглежданията тук посочват възможност да се избегне необходимостта отрицателните примери да бъдат непременно от типа на почти-но-не отрицателни примери, които както вече знаем не се създават лесно

Същност на метода пространството на версиите.

Пространството на версиите може да се изгражда като две допълващи се дървета.

- Едното съдържа върхове, свързани с обобщени модели, а
- другото съдържа върхове, свързани със специализирани модели.

Всяка дъга между върхове в пространството на версиите
представява операция

- специализация или
- обобщаване

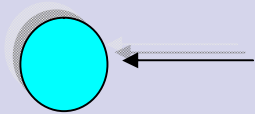
на моделите свързани към върховете.

Ключовата идея на обучението чрез пространството на версиите е, че *специализацията на обобщените модели* и *обобщаването на специфичните модели* в края на краищата води само до един гарантирано правилен модел, който съответства на всички дадени положителни примери и не съответства на нито един от отрицателните примери.

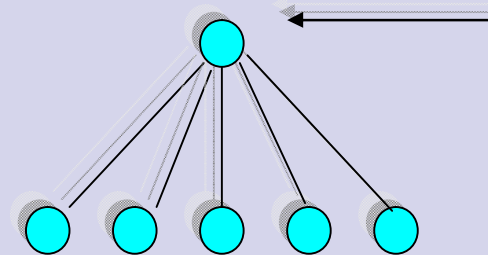
Пространството на версиите е представяне

В което:

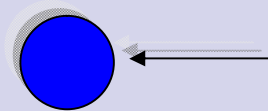
- Има специализиращо дърво и обобщаващо дърво.
- Всеки връх е свързан с модел.
- Обобщаващото дърво започва с корен, който съответствува на възможно най-обобщения модел (той е приложим за всички примери).
- Специализиращото дърво започва с корен, който съответствува само на един положителен пример.
- Дъгите между възлите представят обобщаващите и специализиращите отношения между съответните им модели.



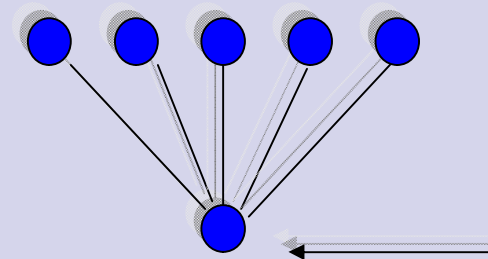
Най-обобщен модел, който съответства на всеки възможен пример



Отрицателните примери специализират обобщените модели.



Най-специализиран модел, който съответства на първия положителен пример.



Положителните примери обобщават специализираните модели.

В началото върховете са само два.

*Единият представя най-обобщения възможен модел, който съответства на всеки възможен обучаващ пример.

**Другият представя специализиран модел, който съответства на описанието на първия представен положителен пример.

Най-обобщеният модел има такива специализации, че нито един от извлечените от него нови модели не съответствуват на първия отрицателен пример. Всеки нов модел е минимално по-различен от модела, от който е извлечен.

Най-специализираният модел има такива обобщения, че всеки от получените от него нови модели са достатъчно общи, за да съответстват на новия положителен пример. Отново всеки от новите модели е малко и уникално изменен.

Така обобщените модели стават по-специализирани, за да намалят съответствията (но без да допускат съответствията с отрицателни примери), а специализираните модели стават по-общи, за да обхванат повече съответствия(но отново без да допускат съответствията с отрицателни примери).

Двете дървета не се разширяват вечно, тъй като много от разклоненията се окастрят.

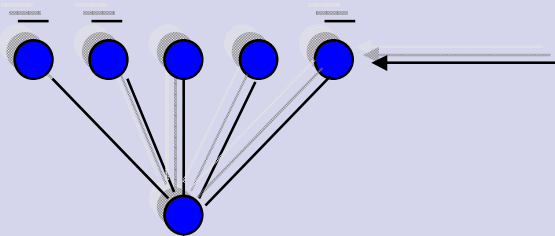
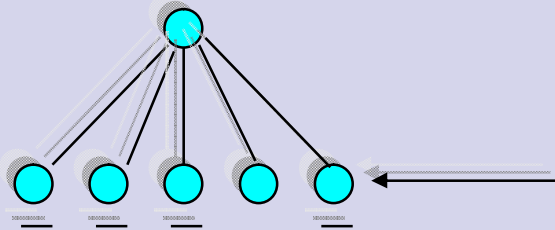
Когато един специализиран модел отговаря (в смисъл не противоречи) на отрицателен пример, този специализиран модел трябва да се елиминира.

Причината е, че специализираните модели се изменят, когато се срещне положителен пример, а самата промяна е в посока на обобщаване. Тогава, ако получен вече специализиран модел обхваща и отрицателен пример, всеки от получените от него специализиран модел ще обхваща този отрицателен пример.

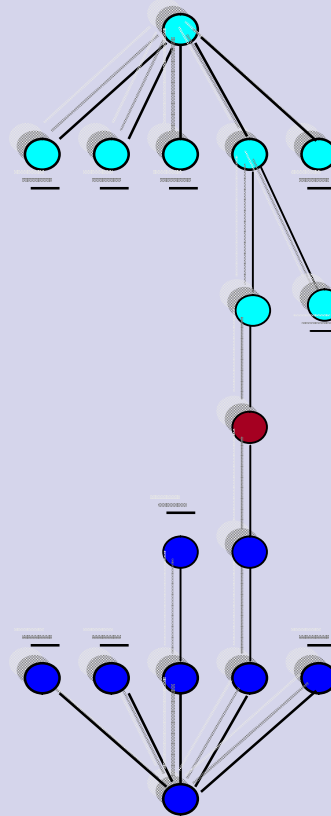
Премахват се и онези от обобщените модели, които грешат при сравняването им с положителни примери. Причината е, че обобщените модели могат само да стават по-специализирани (с по-ограничен обхват). Следователно, ако обобщен модел греша при сравняването му с положителен пример (в смисъл не включва положителния пример), всеки друг обобщен модел извлечен от него ще греша при това сравняване.

Ако приемем че съществува модел решение, положителните и отрицателните примери могат да бъдат само такива, че само един от обобщените модели и само един от специализираните модели ще оцелеят. Тези два модела ще съвпаднат. Когато това стане, повече примери не са необходими, защото е сигурно, че това оцеляло решение е вярно

Положителните примери
окастрият обобщеното описание

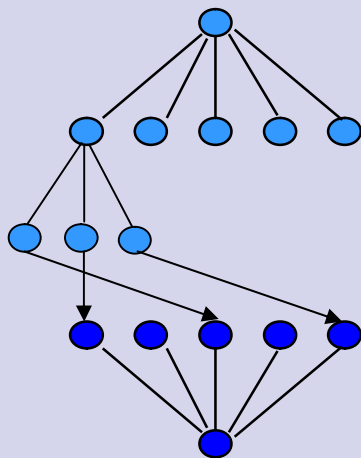


Отрицателните примери
окастрият специализираното описание



**Положителните и отрицателните
примери заставят обобщеното
и специализираното дървета
да се срещнат.**

Какви са допустимите промени в обобщените модели и в специализираните модели?



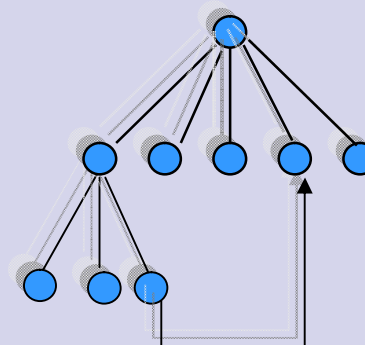
Обобщеният модел се специализира, като тази специализация трябва да бъде обобщение на един от специализираните модели. В противен случай не би имало комбинация от промени, която да доближи специализацията до нещо, което може да се получи от обобщаването на специализираните модели.

Подобно всеки път специализираният модел се обобщава като обобщението трябва да бъде специализация на един от обобщените модели. В противен случай не би имало комбинация от промени, която да доближи новото обобщаване до нещо, което може да се получи от специализацията на обобщените модели.

Всяка специализация на обобщен модел трябва да бъде обобщение на специализиран модел.

Какви са допустимите промени в обобщените модели и в специализираните модели?

Обобщеният модел се специализира така, че тази специализация не трябва да бъде специализация на друг обобщен модел. В противен случай вие бихте могли да запомните ненужен специфичен модел сред обобщените модели.



Специализацията не трябва да бъде специализация на друг обобщен модел.

ОБОБЩЕНИЕТО И СПЕЦИАЛИЗАЦИЯТА ВОДЯТ ДО СХОДИМОСТ НА ПРОСТРАНСТВОТО НА ВЕРСИИТЕ

Като имаме знанията за пространството на версиите в абстрактни термини да разгледаме конкретен пример.

Да предположим, че лекар работи с пациент, който понякога има алергия. Нека интуицията на лекаря му подсказва, че алергията на пациента е резултат от мястото, където се храни, вида хранене, деня от седмицата, цената на храната. Например причина за алергична реакция може да бъде закуска в петък или скъпо ястие в ресторанта на Сам и др.

Така за нашия пример имаме пространство със следните четири координати:

[Ресторант - Хранене - Ден - Цена]

Един конкретен специализиран модел в това пространство е списъкът от конкретни стойности за всяка от координатите , напрмер:
[Ресторанта на Сам, Вечеря, Вторник, Скъпа]

Един изцяло обобщен модел не определя конкретни стойности за атрибутите. Нотация за обобщения модел може да бъде например поставянето ? на мястото на стойността на даден атрибут.

[?, ?, ?, ?]

Въпросителната означава, че е разрешена всяка стойност за съответния атрибут в съответната му позиция.

Между тези два (най-обобщен и най-специализиран) модела са разположени модели, за които някои атрибути са определени, а други не. Например моделът **[На Сам, ?, ?, Скъпа]** съответства на всички случаи, в които пациентът се храни със скъпа храна в ресторанта на Сам по всяко време на деня и в произволен ден от седмицата.

Всеки модел може да се обобщава като на мястото на конкретната стойност за даден атрибут се постави въпросителна.

Всеки модел, който съдържа въпросителна, може да се специализира като въпросителната се замени с конкретна стойност на съответния атрибут.

Да предположим, че примерите по реда на тяхното анализиране са показани в следната таблица:

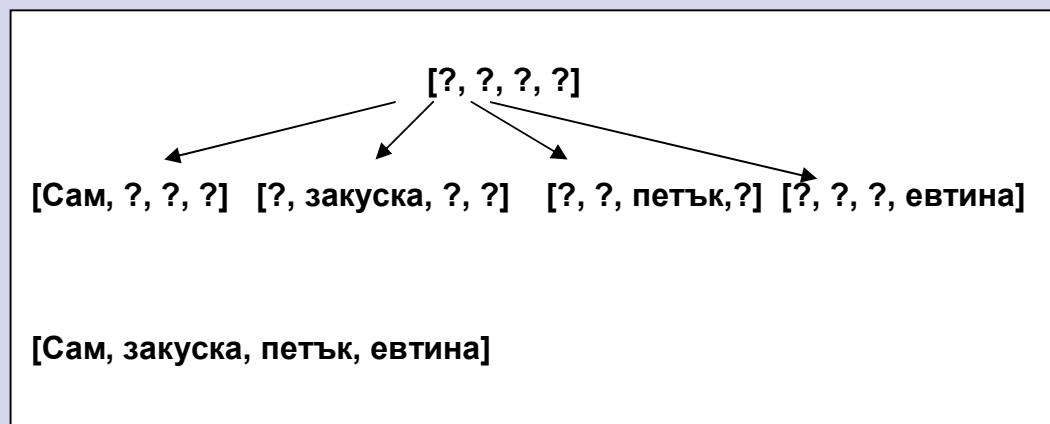
Номер	Ресторант	Хранене	Ден	Цена	Реакция
1	Сам	Закуска	Петък	Евтина	Да
2	Рубин	Обяд	Петък	Скъпа	Не
3	Сам	Обяд	Събота	Евтина	Да
4	Сара	Закуска	Неделя	Евтина	Не
5	Сам	Закуска	Неделя	Скъпа	Не

Първият пример е положителен и той се превръща в първия специализиран модел, като поражда пространството на версиите. То се състои от два модела: най-общия модел и най-специализиран модел

[?, ?, ?, ?]

[Сам, закуска, петък, евтина]

Първият положителен пример специализира обобщения модел. Тъй като се изисква всяка специализация на обобщения модел да е и обобщение на специализирания модел, затова се появяват няколко нови обобщени модела. Забележете, че всяка специализация въвежда минимален брой промени и затова новите обобщени модели са само четири.:



Вторият пример [Рубин, обяд, петък, скъпа] е отрицателен и окастря третото разклонение на обобщените модели, тъй като то не му противоречи

Ако развием всички възможни специализации (за трите дадени ресторанта, за три хранения, за седем дни от седмицата, за вариантите скъпа или евтина храна) ще имаме $3+3+7+2=15$ различни начина на специализиране. 12 от тези 15 специализации грешат при обобщението на специфичния модел.

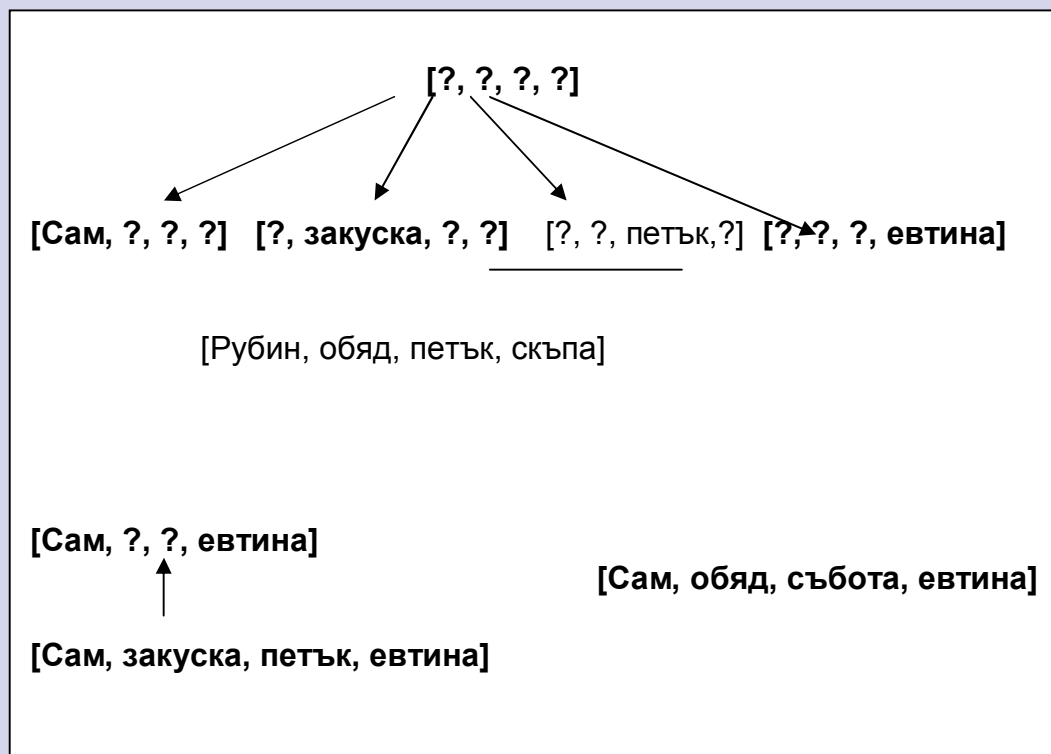


Третият пример- [Сам, обяд, събота, евтина]– е друг положителен пример, който обобщава специализираните модели. Конструираме нов модел като сравняваме съществуващия специализиран модел с положителния пример.

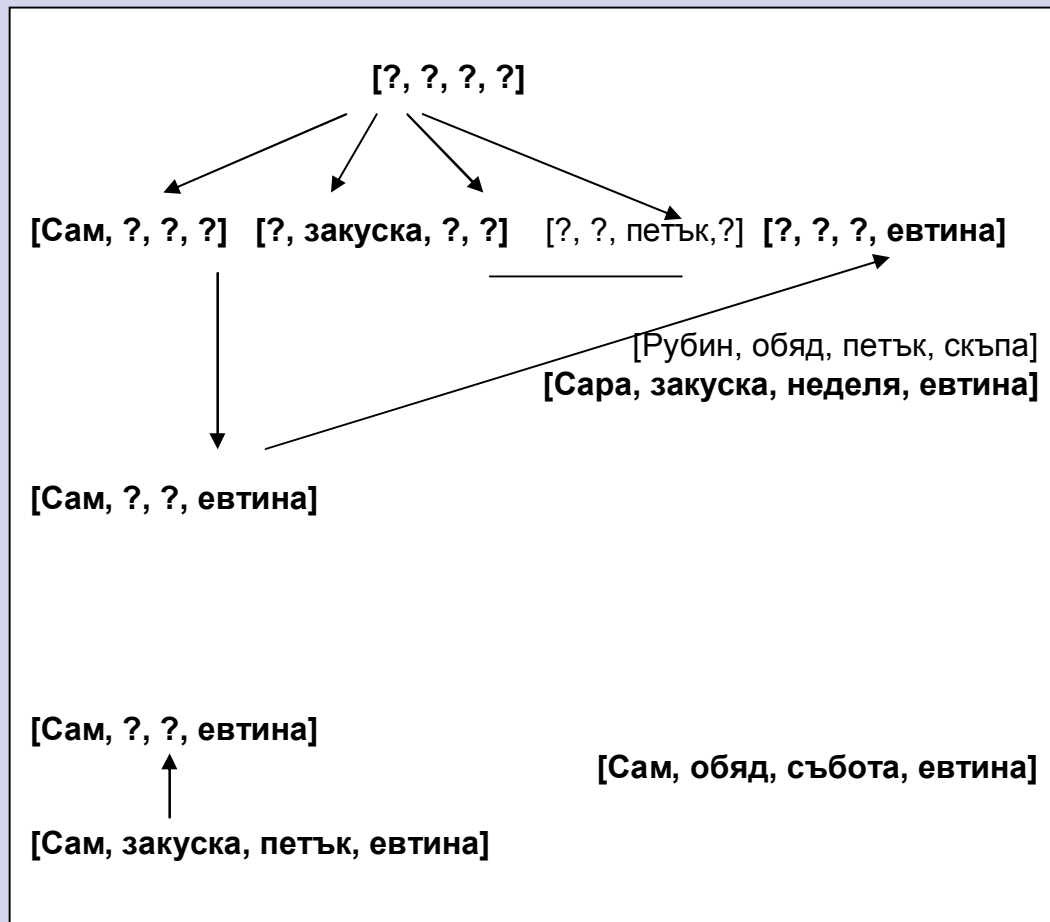
Получава се:

[Сам, ?, ?, евтина].

Този модел обобщава специализирания модел, като забележете че новият модел не обхваща (не противоречи на) отрицателния пример [Рубин, обяд, петък, скъпа].



Четвъртият пример [Сара, закуска, неделя, евтина] е отрицателен и отново окастрия разклонения на обобщените модели (Моделът [?, закуска, ?, ?] не му противоречи и затова не подлежи на развитие).




Специализацията трябва да се съобрази с най-обобщения специализиран модел. Това е моделът [Сам, ?, ?, евтина], който е и единствен.

Ясно е че има само един начин да конкретизираме този модел.

Получаваме моделът [Сам, ?, ?, евтина].

Така този нов обобщен модел става идентичен на специализирания модел

- 
- Когато обобщените модели и специализираните модели се свеждат един към друг, може да сме сигурни, че положителните и отрицателните примери не позволяват друг модел.
 - И така може да се направи заключението, че евтината храна в ресторанта на Сам предизвиква алергична реакция у пациента.

В обобщен вид процедурата може да се опише по следния начин

1) Ако примерът е положителен:

а) Обобщават се всички специализирани модели така, че да съответстват на положителния пример като се осигурява следното:

*** Новите специализирани модели включват минимални промени.**

**** Всеки нов специализиран модел е специализация на даден обобщен модел.**

***** Никой нов специализиран модел не е обобщение на някой друг специализирани модел.**

б) Отстраняват се всички обобщени модели, които не съответстват на положителните примери.

2) Ако примерът е отрицателен,

а) Специализират се всички обобщени модели като не се допуска да съответствуват на отрицателен пример като се осигурява следното:

*** Новите обобщени модели да включва минимални промени.**

**** Всеки нов обобщен модел е обобщение на някой специализиран модел.**

***** Никой обобщен модел не е специализация на някой друг обобщен модел.**

б) Отстраняват се всички специализирани модели, които съответстват на отрицателен пример.

Процедурата води както до множество от специализации, така и до множество обобщения, което не може да се види на нашия твърде елементарен пример.

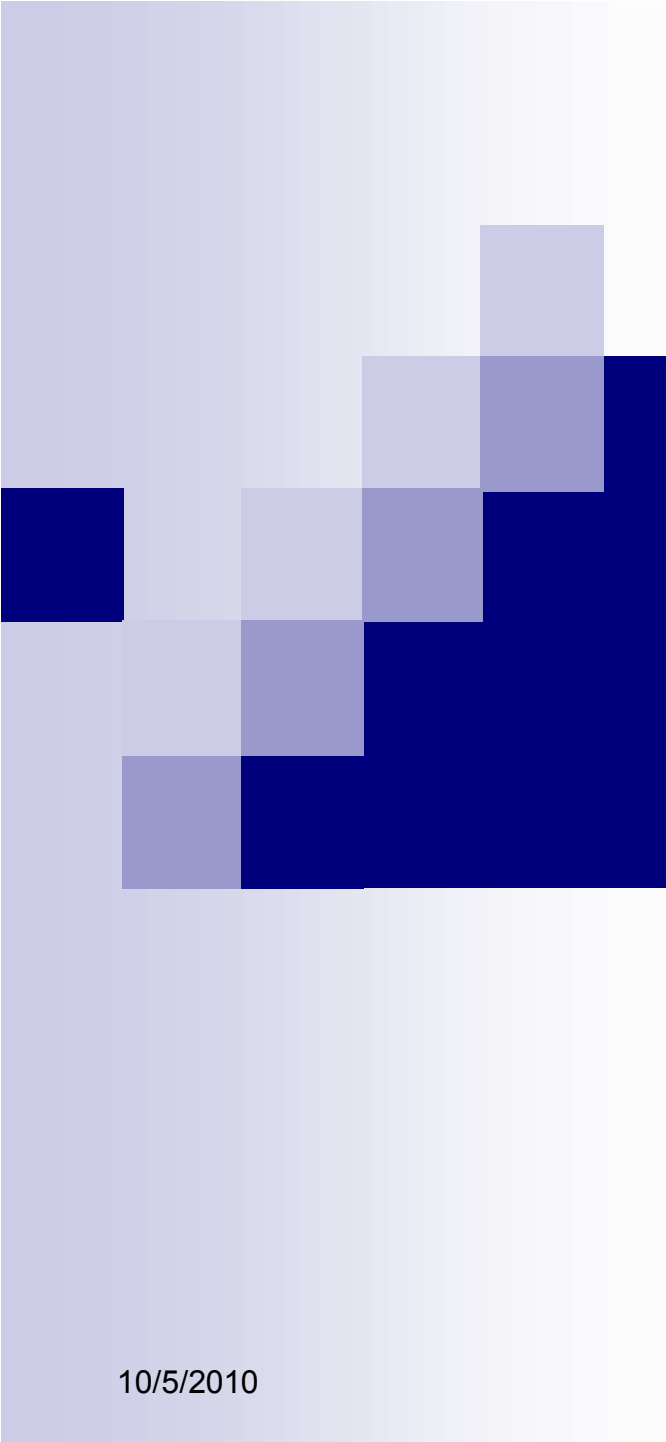
Заключение:

Един начин да учим е да поддържаме множество от прекомерно обобщени модели и множество от прекомерно специализирани модели в пространството на версиите. В такива случаи може да се използват положителни и отрицателни примери, за да се търси сходимост на двете множества до един единствен подходящ модел.

Процедурата на пространството на версиите дава възможност да се постига резултат далеч преди да се получи пълното развитие на обобщеното и на специализираното дървета.

Процедурата на пространството на версиите работи симетрично с положителните и с отрицателните примери. Положителните примери обобщават специализираните модели и ограничават обобщените модели. Отрицателните примери специализират обобщените модели и ограничават специализираните модели.

За разлика от обучението чрез анализ на разликите, където предпазването от погрешни модификации на изграждания модел изисква всеки следващ пример да се отличава съвсем малко от текущия модел при работата в пространството на версиите предпазването от погрешни модификации на изграждания модел става като следването на всяка интерпретация продължава до тогава докато тя е жизнеспособна.



ПОДХОД КЪМ простаряването на версиите



Обучение чрез примери

Зададено


- Правилно класифицирани обучаващи примери

Търси се

- Обобщението, което описва примерите по по-компактен начин

Цел


- Да се класифицират правилно нови неизвестни до сега примери



Сложност на задачата - NP

- n – брой признаци
- k – максимален брой стойности за признак
- Възможните случаи са k^n
- Възможните хипотези са 2^{k^n}

Ако $n = 5$ и $k = 2$ тогава
хипотезите са 2^{32} !!!



Обучение в пространството на версиите

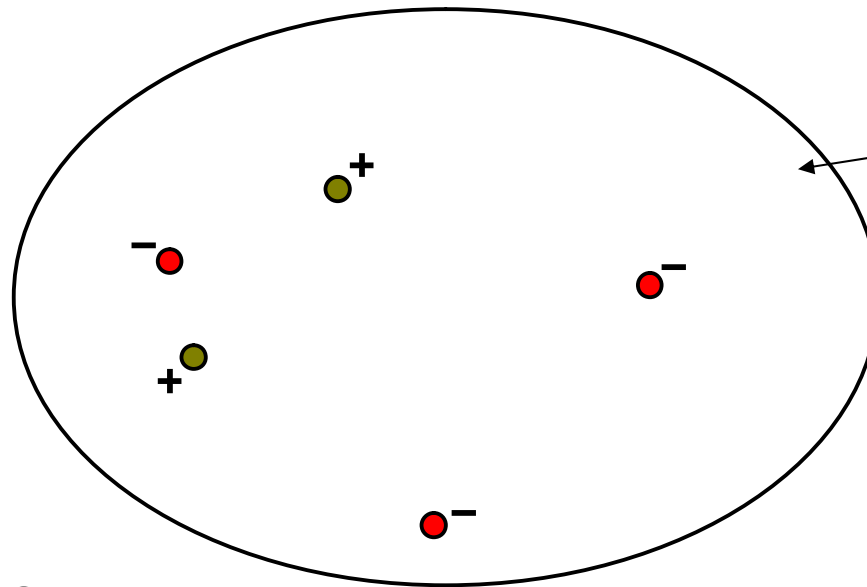
Затруднения:

*Ограничени изразни възможности

**Размерът на общото пространство може да нараства експоненциално от броя на обучаващите примери

***Не е възможно да се обработват зашумени примери или примери с наличие на несигурност

Възможен е следния подход.



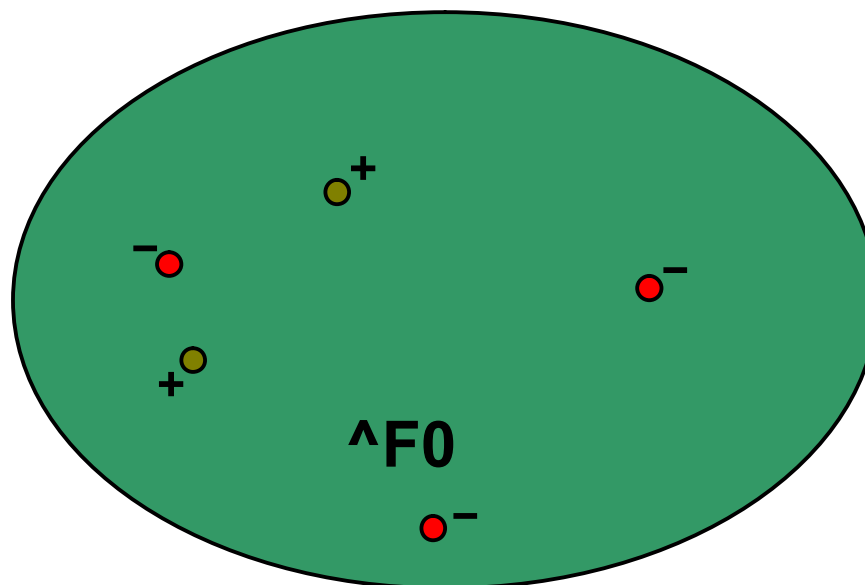
Множество на
ВСИЧКИ
ВЪЗМОЖНИ
случаи

Множество на
известните
положителни
и известните
отрицателни
примери

→ $F1 = P_1 \vee P_2$

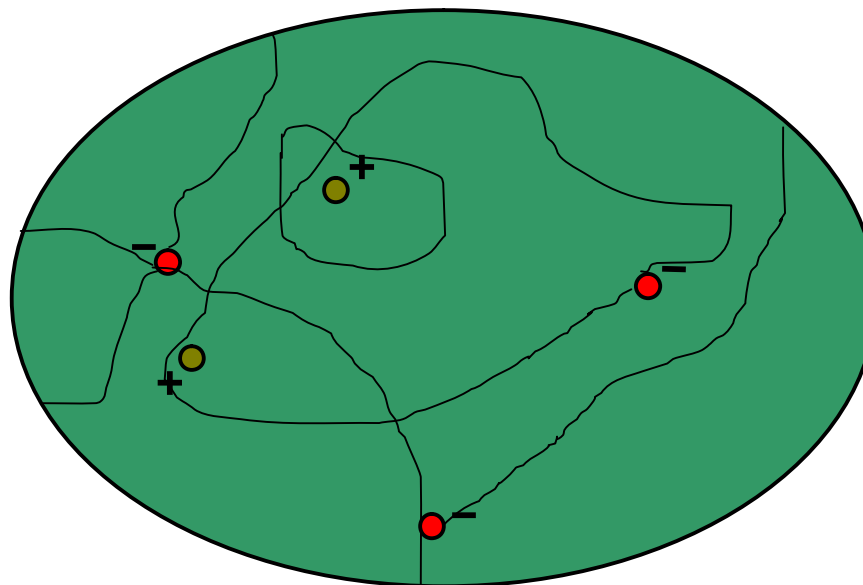
→ $F0 = N_1 \vee N_2 \vee N_3$

СЪЩНОСТ



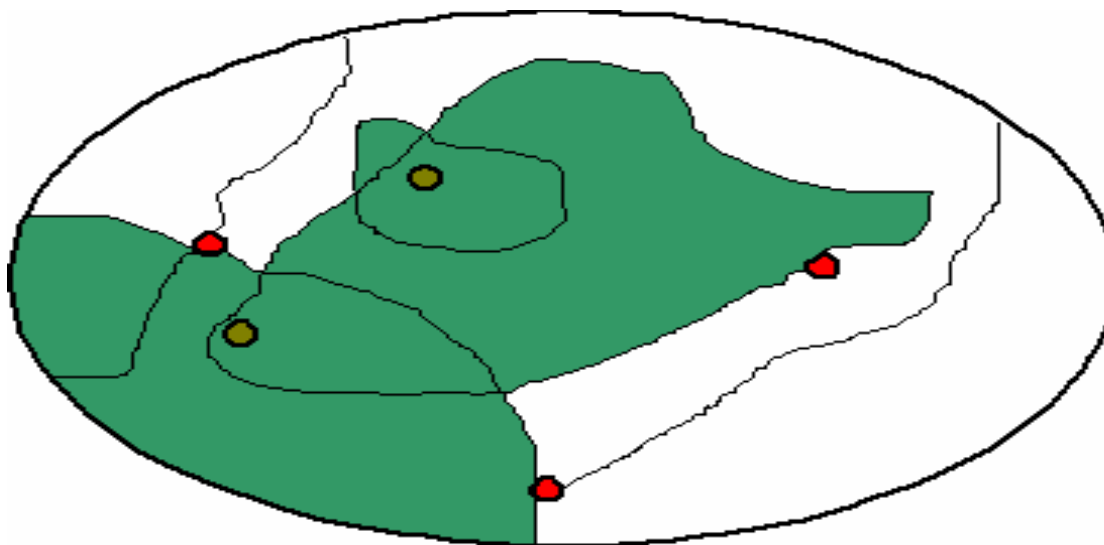
$\wedge F0$ е отрицанието на известните отрицателни примери и би трябвало да покрива всеки от положителните примери

СЪЩНОСТ



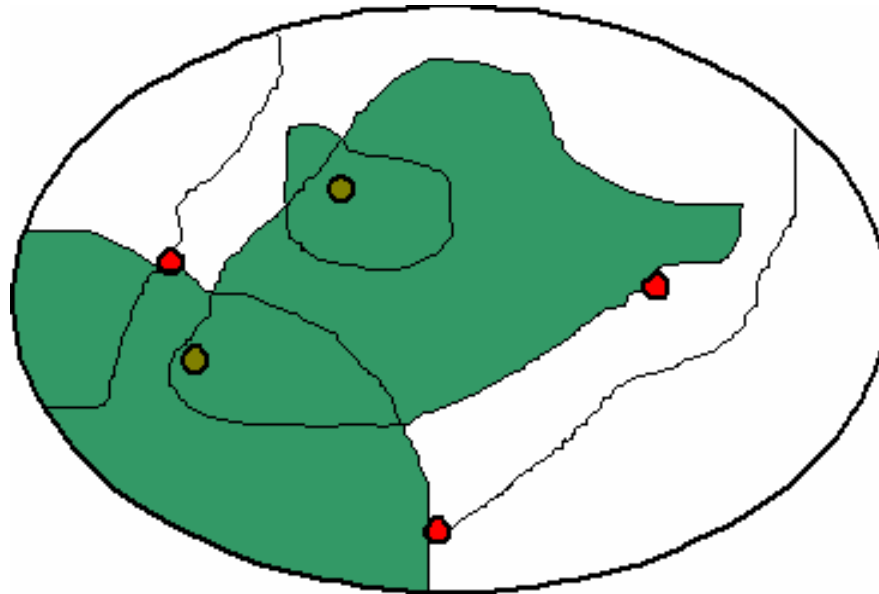
Представянето на $\wedge F_0$ с прости импликанти е едно по-компактно описание

СЪЩНОСТ



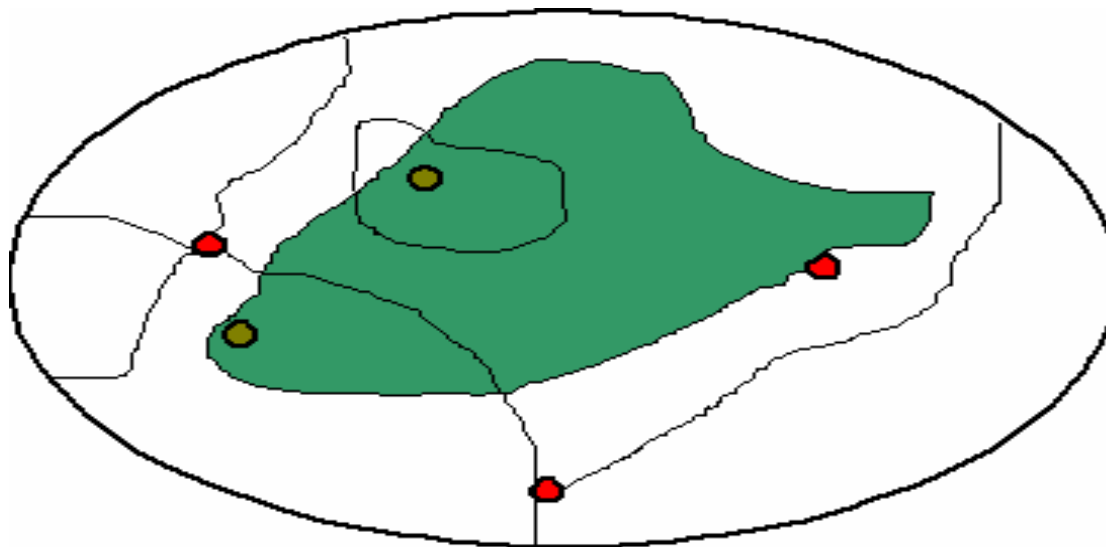
Но на нас са ни необходими само онези импликанти които покриват само известните ни положителни примери

Настройка на представянето на класовете



Фактор на настройката = 1

Настройка на представянето на класовете



Фактор на настройката = 2

Да се върнем към разгледания вече пример

Номер	Ресторант	Ден	Хранене	Цена	Реакция
1	Сам	Петък	Закуска	Евтина	Да
2	Рубин	Петък	Обяд	Скъпа	Не
3	Сам	Събота	Обяд	Евтина	Да
4	Сара	Неделя	Закуска	Евтина	Не
5	Сам	Неделя	Закуска	Скъпа	Не

F1 = Сам * Петък * закуска * евтина
V Сам * Събота * обяд * скъпа

F0 = Рубин * Петък * обяд * скъпа
V Сара * Неделя * закуска * евтина
V Сам * Неделя * закуска * скъпа

Пример

Номер	Ресторант	Ден	Хранене	Цена	Реакция
1	Сам	Петък	Закуска	Евтина	Да
2	Рубин	Петък	Обяд	Скъпа	Не
3	Сам	Събота	Обяд	Евтина	Да
4	Сара	Неделя	Закуска	Евтина	Не
5	Сам	Неделя	Закуска	Скъпа	Не

$\wedge F0 = (\wedge \text{Рубин} \vee \wedge \text{Петък} \vee \wedge \text{обяд} \vee \wedge \text{скъпа})$
* ($\wedge \text{Сара} \vee \wedge \text{Неделя} \vee \wedge \text{закуска} \vee \wedge \text{евтина})$
* ($\wedge \text{Сам} \vee \wedge \text{Неделя} \vee \wedge \text{закуска} \vee \wedge \text{скъпа})$



$\wedge F0 = (\wedge \text{Рубин} \vee \wedge \text{Събота} \vee \wedge \text{обяд} \vee \wedge \text{скъпа})$
* ($\wedge \text{Сара} \vee \wedge \text{Неделя} \vee \wedge \text{закуска} \vee \wedge \text{евтина})$
* ($\wedge \text{Сам} \vee \wedge \text{Неделя} \vee \wedge \text{закуска} \vee \wedge \text{скъпа})$

Или след инвертирането

$\wedge \text{Рубин} = \text{Сара} \vee \text{Сам};$
 $\wedge \text{Петък} = \text{понеделник} \vee \text{вторник} \vee \text{сряда} \vee \text{четвъртък} \vee \text{събота} \vee \text{неделя};$
 $\wedge \text{обяд} = \text{закуска} \vee \text{вечеря};$
 $\wedge \text{скъпа} = \text{евтина}$ и т.н.

$\wedge F0 = ((\text{Сара} \vee \text{Сам}) \vee (\text{понеделник} \vee \text{вторник} \vee \text{сряда} \vee \text{четвъртък} \vee \text{събота} \vee \text{неделя}) \vee (\text{закуска} \vee \text{вечеря}) \vee \text{евтина})$
* ($(\text{Сам} \vee \text{Рубин}) \vee (\text{понеделник} \vee \text{вторник} \vee \text{сряда} \vee \text{четвъртък} \vee \text{петък} \vee \text{събота}) \vee (\text{обяд} \vee \text{вечеря}) \vee \text{скъпа})$
* ($(\text{Сара} \vee \text{Рубин}) \vee (\text{понеделник} \vee \text{вторник} \vee \text{сряда} \vee \text{четвъртък} \vee \text{петък} \vee \text{събота}) \vee (\text{обяд} \vee \text{вечеря}) \vee \text{евтина})$

Сам * Петък * закуска * евтина

$\wedge F0,1 = (\text{Сам} \vee \text{закуска} \vee \text{евтина})$
* ($\text{Сам} \vee \text{Петък})$
* ($\text{Петък} \vee \text{евтина}) = \dots$

Пример

Номер	Ресторант	Ден	Хранене	Цена	Реакция
1	Сам	Петък	Закуска	Евтина	Да
2	Рубин	Петък	Обяд	Скъпа	Не
3	Сам	Събота	Обяд	Евтина	Да
4	Сара	Неделя	Закуска	Евтина	Не
5	Сам	Неделя	Закуска	Скъпа	Не

$$\begin{aligned} \wedge F_{0,1} &= (\text{Сам} \vee \text{закуска} \vee \text{евтина}) \\ &\quad * (\text{Сам} \vee \text{Петък}) \\ &\quad * (\text{Петък} \vee \text{евтина}) = \dots \end{aligned}$$

$$\begin{aligned} \wedge F_{0,2} &= (\text{Сам} \vee \text{Събота} \vee \text{евтина}) \\ &\quad * (\text{Сам} \vee \text{Събота} \vee \text{обяд}) \\ &\quad * (\text{обяд} \vee \text{евтина}) = \dots \end{aligned}$$

Настройка

$$\begin{aligned}
 ^\wedge F_{0,1} &= (\text{Сам} \vee \text{закуска} \vee \text{евтина}) * (\text{Сам} \vee \text{Петък}) * (\text{Петък} \vee \text{евтина}) = \\
 &= (\text{Сам} \vee \text{закуска} \vee \text{евтина}) * (\text{Петък} \vee \text{Сам} * \text{евтина}) = \\
 &= \text{Сам} * \text{Петък} \vee \text{Сам} * \text{евтина} \vee \text{закуска} * \text{Петък} \vee \text{закуска} * \text{Сам} * \text{евтина} \vee \\
 &= \text{евтина} * \text{Петък} \vee \text{Сам} * \text{евтина} \\
 &= \underline{\text{Сам} * \text{Петък} \vee \text{Сам} * \text{евтина} \vee \text{Петък} * \text{закуска} \vee \text{Петък} * \text{евтина}}
 \end{aligned}$$

$$\begin{aligned}
 ^\wedge F_{0,2} &= (\text{Сам} \vee \text{Събота} \vee \text{евтина}) * (\text{Сам} \vee \text{Събота} \vee \text{обяд}) * (\text{обяд} \vee \text{евтина}) = \\
 &= (\text{Сам} \vee \text{Събота}) * (\text{обяд} \vee \text{евтина}) * (\text{обяд} \vee \text{евтина}) = \\
 &= \underline{\text{Сам} * \text{обяд} \vee \text{Сам} * \text{евтина} \vee \text{Събота} * \text{обяд} \vee \text{Събота} * \text{евтина}}
 \end{aligned}$$

Импликанта	Сам* Петък	Сам* евтина	Петък * закуска	Петък * евтина	Сам* обяд	Събота* обяд	Събота* евтина
Фактор настройка	1	2	1	1	1	1	1



Заклучение

- Удобен алгоритъм
- Решава задачата по-добре от останалите методи
- Висока точност
- Позволява настройка



Възможно развитие

- Приложим е и към цифрови и структурирани видове на параметрите
- Решенията могат да се изобразят като дърво на решенията, от което да се генерират правила
- Приложим е и за множество от класове
- Притежава добри възможности за използване при извличане на данни



Литература.

[Mitchell, T.M., Generalization as Surch, *Artificial Intelligence*, vol 18, no 2, 1982] .

[Dakovski L., Zekie Shevked, (2005): Alternative Approach for Learning from Examples, CompSysTech'05, Proceedings, pp.III.B.5-1-IIIB.5-6, ISBN 954-9641-42-2, Varna.]