

Introduction to Data and Knowledge Engineering Sommersemester 2010



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Übung 9 Separate and Conquer

25. Juni 2010

Aufgabe 9.1 Separate and Conquer

In dieser Übung soll ein Separate-and-Conquer Algorithmus durchgeführt werden, der die folgenden Fakten als Hintergrundwissen verwendet.

```
philanthrop(bertha)
philanthrop(peter)
```

```
male(dieter)
male(peter)
```

```
employee(anna, dieter)
employee(bertha, peter)
employee(carla, anna)
employee(dieter, peter)
employee(peter, anna)
```

Das zu trainierende Prädikat ist likes/2 und die folgenden Beispiele stehen für das Training zur Verfügung.

```
⊕: likes(bertha, carla)   ⊖: likes(anna, anna)
   likes(bertha, peter)  likes(carla, anna)
   likes(dieter, bertha) likes(dieter, peter)
   likes(dieter, anna)   likes(anna, carla)
   likes(peter, carla)
   likes(peter, dieter)
```

Als Heuristik zur Auswahl des nächsten Literals soll anstelle des von FOIL verwendeten Weighted Information Gains die Differenz $p - n$ herangezogen werden, wobei p die Anzahl der positiven und n der Anzahl negativen verbleibenden Beispiele ist, welche von dem Literal abdeckt werden.

Desweiteren wollen wir zwei Vereinfachungen machen. Erstens sei die Menge der Kandidaten-Literale wie in der unten stehenden Tabelle vorgegeben. Einige Kandidaten, die sonst ebenfalls geprüft würden, wurden hier weggelassen. Zweitens wollen wir nach der zweiten gefundenen Regel abbrechen, unabhängig davon, ob dann alle positiven Beispiele aus der Trainingsmenge abgedeckt sind. Die ersten beiden freien Spalten der Tabelle sind für die erste Regel vorgesehen, die anderen für die zweite Regel.

	p	n	$p - n$	p	n	$p - n$	p	n	$p - n$	p	n	$p - n$
male(X)												
not male(X)												
male(Y)												
not male(Y)												
philanthrop(X)												
philanthrop(Y)												
employee(X, Y)												
employee(Y, X)												
employee(X, Z)												