

3. Übungsblatt

Aufgabe 10 Apriori-Algorithmus

- a) Führen Sie den Apriori-Algorithmus (nach Pseudo-Kode) für die rechts gezeigte Transaktionsdatenbank und einen minimalen Support von $s_{\min} = 3$ aus!
- | | |
|-----------|-----------|
| $a d f$ | $a b d$ |
| $a c d e$ | $b d e$ |
| $b d$ | $b c e g$ |
- b) Führen Sie den Apriori-Algorithmus unter Verwendung eines Präfixbaumes zur Organisation der Zähler aus!
- | | |
|---------|---------|
| $b c d$ | $c d f$ |
| $b c$ | $a b d$ |
- c) Wie wird der Support im Apriori-Algorithmus bestimmt (Pseudo-Kode)?
Was sind die Vor- und Nachteile dieses Zählschemas?
Wie wird der Support bei Organisation der Zähler als Präfixbaum bestimmt?
(Hinweis: Vergleiche Teilmengentests und Erzeugen und Finden von Teilmengen.)
- d) Was sind die Vorteile der Organisation der Transaktionen als Präfixbaum?
(Hinweis: Welche Redundanz kann durch diesen Ansatz vermieden werden?)
(Achtung: Verwechseln Sie nicht den Präfixbaum zur Organisation der Supportzähler mit dem Präfixbaum zur Darstellung der Transaktionen!)

Aufgabe 11 Teile-und-Herrsche-Suche

- a) Wie kann die Suche nach häufigen Itemmengen als Teile-und-Herrsche-Verfahren beschrieben werden? Wie kommt man so zu einem rekursiven Suchverfahren?
- b) Wie kann das Teile-und-Herrsche-Verfahren formal als Menge von Teilproblemen beschrieben werden? Wie wird jedes Teilproblem beschrieben?
- c) Wie wird ein Teilproblem gelöst? Insbesondere, wie werden die Elemente der Teilprobleme bestimmt, in die das zu lösende Teilproblem aufgeteilt wird?
- d) Warum sind die Transaktionsdatenbanken, die zur Lösung der beiden Teilprobleme benötigt werden, *keine* Komplemente voneinander?

Aufgabe 12 Perfekte Erweiterungen

- a) Was ist eine perfekte Erweiterung? Wie wird sie formal definiert?
- b) Seien a und b perfekte Erweiterungen einer Itemmenge I und sei $c \notin I$ ein weiteres Item (keine perfekte Erweiterung).
Ist a perfekte Erweiterung von $I \cup \{b\}$? Ist a perfekte Erweiterung von $I \cup \{c\}$?
Ist b perfekte Erweiterung von $I \cup \{a\}$? Ist b perfekte Erweiterung von $I \cup \{c\}$?
Was kann man allgemein aus den Antworten auf die obigen Fragen ablesen?
- c) Falls im Teile-und-Herrsche-Verfahren das Aufteilungsitem eine perfekte Erweiterung ist, wie verhalten sich die Lösungen der beiden Teilprobleme zueinander? Warum? (Hinweis: Vergleichen Sie die zugehörigen Transaktionsdatenbanken.)
- d) Wie kann man mit perfekten Erweiterungen das Teile-und-Herrsche-Suche vereinfachen? (Hinweis: Müssen immer beide Teilprobleme gelöst werden?)

Aufgabe 13 Darstellung von Transaktionsdatenbanken

- a) Bestimmen Sie horizontale und vertikale Darstellungen der Matrix

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
1:	1	0	0	1	0	1	0
2:	1	0	1	1	1	0	0
3:	0	1	0	1	0	0	0
4:	0	1	1	1	0	0	0
5:	0	1	1	0	0	0	0
6:	1	1	0	1	0	0	0
7:	0	1	0	1	1	0	0
8:	0	1	1	0	1	0	1
9:	0	0	1	1	0	1	0
10:	1	1	0	1	0	0	0

- b) Kann man den Apriori-Algorithmus mit einer vertikalen Darstellung der Transaktionsdatenbank ausführen?
- c) Kann man eine horizontale Darstellung der Transaktionsdatenbank im Teile-und-Herrsche-Verfahren zum Finden häufiger Itemmengen verwenden?
- d) Charakterisieren Sie eine Präfixbaumdarstellung einer Transaktionsdatenbank in Begriffen horizontaler/vertikaler Darstellungen!