

Übung Informationsintegration

Schema Matching

Aufgabe 1 (Stable Marriage):

Gegeben ist folgende Ähnlichkeitsmatrix:

	A	B	C	D
1	0.8	0.2	0.1	0.9
2	0.6	0.0	0.8	0.7
3	0.5	0.3	0.6	0.4
4	0.4	0.1	0.5	0.6

- Bestimmen Sie die Präferenzlisten der Herren (1-4) und der Damen (A-D)
- Führen Sie den in der Vorlesung vorgestellten Algorithmus zum Finden einer stabilen Ehe aus Sicht der Herren aus (z.B. "1 stellt Antrag an A, sie ...")
- Führen Sie den selben Algorithmus aus Sicht der Damen aus (z.B. "A stellt Antrag an 1, er ...")

Aufgabe 2 (A*-Suche):

Gegeben sind zwei Schemata \mathcal{A} und \mathcal{B} mit je drei Elementen $\{A_1, A_2, A_3\}$ bzw. $\{B_1, B_2, B_3\}$. Die Ähnlichkeitsmatrix dieser Elemente ist:

	B_1	B_2	B_3
A_1	1	1	1
A_2	1	0	1
A_3	0	1	1

Im folgenden wird eine Match Combination in der Form $(B_1, B_2, *)$ dargestellt, wobei das Symbol $*$ eine Wildcard darstellt.

Zudem sind die folgenden vier Integritätsbedingungen gegeben:

Bedingung	Kosten	Verletzt durch
IC_1	∞	$\{(B_3, B_1, *)\}$
IC_2	2	$\{(B_2, *, *), (*, B_1, B_3)\}$
IC_3	4	$\{(B_1, *, *), (*, B_3, *)\}$
IC_4	1	$\{(B_2, *, *), (B_3, *, *)\}$

Führen Sie eine A*-Suche durch um die 'beste' Match Combination zu finden. Dabei gelten folgende Umstände:

- Als Kosten werden nur diejenigen berücksichtigt, die durch die Integritätsbedingungen verursacht werden (die Güte der Match Combinations wird nicht beachtet)
- Expandieren Sie die Elemente des ersten Schemas der Reihe nach (d.h. zuerst A_1 , dann A_2 und zuletzt A_3)
- Betrachten Sie dabei nur konsistente Match Combinations (Hinweis: Eine Match Combination ist inkonsistent, wenn zwei Elemente des selben Schemas mit dem selben Element des anderen Schemas korrespondieren)
- Match Combinations die eine Korrespondenz enthalten welche keine Glaubhaftigkeit hat (Ähnlichkeit 0) werden von vorne herein ausgeschlossen
- Jedes Schemaelement muss genau eine Korrespondenz eingehen (d.h. der Wert \perp ist nicht erlaubt)