Übungen aus der extra Sitzung: Übungsblatt 4/Aufgabe 3 ; Übungsblatt 3/Aufgabe 6

Übungsblatt 4 / Aufgabe 3: **[ihr könnt direkt zur Lösung überspringen]**

Alles was in dieser Aufgabe benötigt wird, ist halt den Kandidateneliminationsalgorithmus für die Menge der angegebenen 5 Instanzen (d.h. Belegungen, Elementen vom Objektraum, Beispielsdaten) – der ganze Vorgang lässt sich folgendermaßen beschreiben:

Für jeden Schritt, d.h. für jedes Beispiel nacheinander (dabei wird aufgefordert, die Reihenfolge „von oben nach unten" in unserer Datentabelle zu behalten), 2 Mengen S und G zu verfeinern, die am Anfang folgendermaßen definiert werden:

$S = \{(\emptyset, \emptyset, \emptyset, \emptyset, \emptyset)\}$ - die Menge der speziellsten Hypothesen für unseren zu lernenden Begriff (hier - „Example Type"); diese Menge dient als ***die unterste Schranke***, ab der die Hypothesen konsistent mit der angegebenen 5 Lerndaten sind. Also am Anfang enthält S nur die kleinste, nie erfüllbare Hypothese.

$G = \{(?\,,?\,,?\,,?\,,?\,)\}$ - die Menge der generellsten Hypothesen für unseren zu lernenden Begriff (hier - „Example Type"); diese Menge dient als ***die oberste Schranke***, bis der die Hypothesen konsistent mit der angegebenen 5 Lerndaten sind. Also am Anfang besteht G aus der größten Hypothese.

Begriffe „ab" und „bis" zeigen auf die Existenz von einer Relation zwischen Hypothesen, d.h. in diesem Fall auf eine Halbordnung, die den Vergleich zwischen 2 Hypothesen ermöglicht (natürlich nicht alle Hypothesen sind vergleichbar, daher reden wir über die **Halb**ordnung und nicht über die totale Ordnung) – diese Relation ist halt die Inklusion dieser Art:

Wir sagen dass eine Hypothese h_1 genereller als eine andere Hypothese h_2 ist (bzw. h_2 spezieller als h_1 ist) und drücken das $h_1 \supseteq h_2$ ($\supset$ bedeutet $\supseteq$ und $\neq$) aus, genau dann, wenn jede Belegung, die h_2 erfüllt, auch h_1 erfüllt.

Und eine Belegung x erfüllt eine Hypothese h genau dann, wenn jede Komponente von h entweder ? ist, oder gleich der entsprechenden Komponente von x ist ($x \in Origin \times Manufacturer \times Color \times Decade \times Type$ und $h \in Origin \cup \{?\} \times Manufacturer \cup \{?\} \times Color \cup \{?\} \times Decade \cup \{?\} \times Type \cup \{?\}$).

Nun sind wir zum Punkt gekommen, nach der kurzen Beschreibung des durchzuführenden Vorgangs, die Aufgabe in 5 Schritten (jeweils mit einem gleichen Vorgang je nachdem wir ein Positivbeispiel oder ein Negativbeispiel betrachten) direkt zu lösen:

Wenn wir ein Positivbeispiel (+) x an der Reihe haben, kürzen wir (wenn nötig) sofort die Menge G vom vorherigen Schritt (und machen sie zur aktuellen Menge) und arbeiten an der Menge S in 3 Stufen:

- **PG**: „Kürzen" heißt hier die umgehende Elimination von denjenigen Hypothesen aus der Menge G, die vom aktuellen Beispiel **nicht erfüllt** werden;
- **PS**: „3-stufige Arbeit" bedeutet:
  1. *Erstens[**Kernsatz**: ziehe die untere Schranke S **Schritt für Schritt** hoch bis sie komplett zutreffend für x wird]:* für jedes Element h aus S, wenn x diese Hypothese h nicht erfüllt, generalisiert man h zu einer neuen Hypothese h'(d.h. wandelt h in eine neue Hypothese h' um, die genereller als h ist), die von x erfüllt wird. ***Schritt für Schritt*** = „wird jeweils nur ein nichtzutreffender Attributwert x_i (beim ersten Schritt - $\emptyset$) zur Komponente ? (beim ersten Schritt - zur zutreffenden x_i)".
  2. *Zweitens[**Kernsatz**: stelle sicher, dass S spezieller als G bleibt]:* jedes Element h in S, das nicht *spezieller* als G (d.h. als jedes Element aus G) ist, wird aus S gelöscht.
  3. *Drittens[**Kernsatz**: stelle sicher, dass wenn S die Speziellste bleibt]:* keine Hypothese aus S darf *genereller* als eine andere Hypothese aus S sein – wenn $h_1 \supseteq h_2$ ($h_1, h_2 \in S$), dann wird h_1 aus S gelöscht.

Wenn wir ein Negativbeispiel (-) x betrachten, machen wir quasi alles umgekehrt: wir kürzen (wenn nötig) sofort die Menge S vom vorherigen Schritt (und machen sie zur aktuellen Menge) und arbeiten in 3 Stufen an der Menge G:

- **NS**: „Kürzen" heißt hier die umgehende Elimination von denjenigen Hypothesen aus der Menge S, die vom aktuellen Beispiel **erfüllt** werden;
- **NG**: „3-stufige Arbeit" bedeutet:
  1. *Erstens[**Kernsatz**: ziehe die obere Schranke G **Schritt für Schritt** herunter bis sie komplett nichtzutreffend für x wird]:* für jedes Element h aus G, wenn x diese Hypothese h erfüllt, spezifiziert man h zu einer neuen Hypothese h'(d.h. wandelt h in eine neue Hypothese h' um, die spezieller als h ist), die von x nicht erfüllt wird. ***Schritt für Schritt*** = „wird jeweils nur eine Komponente ? zum nichtzutreffenden Attributwert x_i".
     Anmerkung: *dabei konstruiert man unter jede zu betrachtende Komponente jeweils so viele speziellere Hypothesen, wie die Anzahl der bis diesem Punkt aufgetretenen nichtzutreffenden Attributwerte.*
     Zum Beispiel: *beim 2. Schritt, d.h. beim zweiten Beispiel sind unter dem Attribut „**Manufacturer**" nur die Werte „**Honda**" und „**Toyota**" aufgetreten und daher hat man nur einen nichtzutreffenden Attributwert „**Honda**" um die Hypothesen in dieser Komponente zu spezifizieren.*
     Im Prinzip darf man auch die anderen Werte verwenden, allerdings wird damit nur ein (im Rahmen des aktuellen Algorithmus) nutzloser Kram zur Berechnung beigefügt (und kurz danach eliminiert), was alles anderes als effektiv ist ☺.
  2. *Zweitens[**Kernsatz**: stelle sicher, dass G genereller als S bleibt]:* jedes Element h in G, das nicht *genereller* als S (d.h. als jedes Element aus S) ist, wird aus G gelöscht.
     Anmerkung: dieser Punkt verfeinert die im obigen Vorgang spezifizierte Menge G, damit die mit vorherigen Beispielen inkonsistenten Hypothesen entfernt werden können.

3. *Drittens[**Kernsatz**: stelle sicher, dass wenn G die Generellste bleibt]:* keine Hypothese aus G darf *spezieller* als eine andere Hypothese aus G sein – wenn $\mathrm{h}_1 \supseteq \mathrm{h}_2$ $(h_1, h_2 \in G)$, dann wird h_2 aus G gelöscht.

Zum Schluss fassen wir unsere Mengen in einem Mengenintervall $[S; G]$ zusammen, das aus einem Element (meist der beste Fall), aus mehreren Elementen oder aus gar keinem Element (d.h. die leere Menge) bestehen kann.

Dieses Intervall lässt sich folgendermaßen intuitiv interpretieren: all die Hypothesen, die dazwischen fallen (also all die Hypothesen aus S, alle aus G und alle dazwischen), sind mit den angegebenen Lerndaten (in unserem Fall – mit 5 Beispielsdaten) konsistent.

Wenn wir beispielsweise nur ein einziges Element im oben genannten Intervall bekommen, sagt man dass der Begriff (induktiv) gelernt wurde – natürlich gibt's dazu viel mehr Einzelheiten und das ganze Begriffslernen sieht nicht so einfach aus (mehr dazu wird u.a. im Vorlesungsskript erwähnt).

Und jetzt geht es mit der tatsächlichen **Lösung** los:

I Beispiel) $x = (x_1, x_2, x_3, x_4, x_5) = (Japan, Honda, Blue, 1980, Economy)$ - Positivbeispiel

Aktuelle Menge S: $\{(\emptyset, \emptyset, \emptyset, \emptyset, \emptyset)\}$
Aktuelle Menge G: $\{(?, ?, ?, ?, ?)\}$

Teilschritt PG: $G = \{(?, ?, ?, ?, ?)\}$ **– wird zur aktuellen $\boldsymbol{G}$**
Teilschritt PS1: $S = \{(Japan, Honda, Blue, 1980, Economy)\}$
Teilschritt PS2: $S = \{(Japan, Honda, Blue, 1980, Economy)\}$
Teilschritt PS3: $S = \{(Japan, Honda, Blue, 1980, Economy)\}$ **– wird zur aktuellen $\boldsymbol{S}$**

II Beispiel) $x = (x_1, x_2, x_3, x_4, x_5) = (Japan, Toyota, Green, 1970, Sports\ )$ - Negativbeispiel

Aktuelle Menge S: $\{(Japan, Honda, Blue, 1980, Economy\ )\}$
Aktuelle Menge G: $\{(?\,,?\,,?\,,?\,,?\,)\}$

Teilschritt NS: $S = \{(Japan, Honda, Blue, 1980, Economy\ )\}$ **– wird zur aktuellen $\boldsymbol{S}$**
Teilschritt NG1: $G = \{(?\,, Honda, ?\,, ?\,, ?\,), (?\,, ?\,, Blue, ?\,, ?\,), (?\,, ?\,, ?\,, 1980, ?\,), (?\,, ?\,, ?\,, ?\,, Economy)\}$
Teilschritt NG2: $G = \{(?\,, Honda, ?\,, ?\,, ?\,), (?\,, ?\,, Blue, ?\,, ?\,), (?\,, ?\,, ?\,, 1980, ?\,), (?\,, ?\,, ?\,, ?\,, Economy)\}$
Teilschritt NG3: $G = \{(?\,, Honda, ?\,, ?\,, ?\,), (?\,, ?\,, Blue, ?\,, ?\,), (?\,, ?\,, ?\,, 1980, ?\,), (?\,, ?\,, ?\,, ?\,, Economy)\}$ **– wird zur aktuellen $\boldsymbol{G}$**

III Beispiel) $x = (x_1, x_2, x_3, x_4, x_5) = (Japan, Toyota, Blue, 1990, Economy\ )$ - Positivbeispiel

Aktuelle Menge S: $\{(Japan, Honda, Blue, 1980, Economy\ )\}$
Aktuelle Menge G: $\{(?\,, Honda, ?\,, ?\,, ?\ ), (?\,, ?\,, Blue, ?\,, ?\ ), (?\,, ?\,, ?\,, 1980, ?\ ), (?\,, ?\,, ?\,, ?\,, Economy)\}$

Teilschritt PG: $G = \{(?\,, ?\,, Blue, ?\,, ?\ ), (?\,, ?\,, ?\,, ?\,, Economy)\}$ **– wird zur aktuellen *G***
Teilschritt PS1: $S = \{(Japan, ?\,, Blue, ?\,, Economy\ )\}$
Teilschritt PS2: $S = \{(Japan, ?\,, Blue, ?\,, Economy\ )\}$
Teilschritt PS3: $S = \{(Japan, ?\,, Blue, ?\,, Economy\ )\}$ **– wird zur aktuellen *S***

IV Beispiel) $x = (x_1, x_2, x_3, x_4, x_5) = (USA, Chrysler, Red, 1980, Economy)$ - Negativbeispiel

Aktuelle Menge S: $\{(Japan, ?, Blue, ?, Economy)\}$
Aktuelle Menge G: $\{(?, ?, Blue, ?, ?), (?, ?, ?, ?, Economy)\}$

Teilschritt NS: $S = \{(Japan, ?, Blue, ?, Economy)\}$ **– wird zur aktuellen *S***

Teilschritt NG1: $G = \left\{ \begin{matrix} (?, ?, Blue, ?, ?), (Japan, ?, ?, ?, Economy), (?, Honda, ?, ?, Economy), (?, Toyota, ?, ?, Economy), \\ (?, ?, Blue, ?, Economy), (?, ?, Green, ?, Economy), (?, ?, ?, 1970, Economy), (?, ?, ?, 1990, Economy) \end{matrix} \right\}$

Teilschritt NG2: $G = \left\{ \begin{matrix} (?, ?, Blue, ?, ?), (Japan, ?, ?, ?, Economy), \\ (?, ?, Blue, ?, Economy) \end{matrix} \right\}$

Teilschritt NG3: $G = \left\{ (?, ?, Blue, ?, ?), (Japan, ?, ?, ?, Economy) \right\}$ **– wird zur aktuellen *G***

V Beispiel) $x = (x_1, x_2, x_3, x_4, x_5) = (Japan, Honda, White, 1980, Economy\ )$ - Positivbeispiel

Aktuelle Menge S: $\{(Japan, ?\,, Blue, ?\,, Economy\ )\}$
Aktuelle Menge G: $\{(?\,, ?\,, Blue, ?\,, ?\ ), (Japan, ?\,, ?\,, ?\,, Economy)\}$

Teilschritt PG: $G = \ \{(Japan, ?\,, ?\,, ?\,, Economy)\}$ **– wird zur letzten $\boldsymbol{G}$**
Teilschritt PS1: $S = \{(Japan, ?\,, ?\,, ?\,, Economy\ )\}$
Teilschritt PS2: $S = \{(Japan, ?\,, ?\,, ?\,, Economy\ )\}$
Teilschritt PS3: $S = \{(Japan, ?\,, ?\,, ?\,, Economy\ )\}$ **– wird zur letzten $\boldsymbol{S}$**

Übungsblatt 3 / Aufgabe 6:

$$F = \forall z \exists y \big(p(x, g(y), z) \ \lor \ \underbrace{\neg \forall x q(x)}_{G}\big) \quad \land \quad \underbrace{\neg \forall z \exists x \neg r(f(x,z),z)}_{H}$$

$$F = \forall z \exists y \big(p(x, g(y), z) \ \lor \ G \ \big) \ \land \ H$$

1) Umformung von F in eine äquivalente Formel in ***Pränex-Normalform*** (*PNF*):

$$\begin{aligned}
F &\equiv \forall z \exists y \big(p(x,g(y),z) \lor \exists x \neg q(x)\big) \land \exists z \neg \exists x \neg r(f(x,z),z) && \text{Negation nach innen schieben}\\
&\equiv \forall z \exists y \big(p(x,g(y),z) \lor \exists x \neg q(x)\big) \land \exists z \forall x \neg\neg r(f(x,z),z) && \text{Negation nach innen schieben}\\
&\equiv \forall z \exists y \big(p(x,g(y),z) \lor \exists x \neg q(x)\big) \land \exists z \forall x r(f(x,z),z) && \text{Doppelnegation}\\
&\equiv \forall z \exists y \big(p(x,g(y),z) \lor \exists t \neg q(t)\big) \land \exists u \forall v r(f(v,u),u) && \\
&\equiv \forall z \exists y \exists t \exists u \forall v \big((p(x,g(y),z) \lor \neg q(t)) \land r(f(v,u),u)\big) && \text{Quantoren nach außen ziehen}
\end{aligned}$$

Umbennen der gebundenen Variablen (d.h. **Ersetzung** durch neue Variablen): **in** $\boldsymbol{G}$: x/t, **in H**: z/u und x/v

Pränex-Normalform von F:

$$\forall z \exists y \exists t \exists u \forall v \big((p(x,g(y),z) \lor \neg q(t)) \land r(f(v,u),u)\big)$$

2) Weitere Umformung in eine äquivalente Formel in ***Skolem-Normalform*** (*SNF*):

$$\equiv \forall x \forall z \exists y \exists t \exists u \forall v \big((p(x, g(y), z) \vee \neg q(t)) \wedge r(f(v, u), u)\big)$$

$$\equiv \forall x \forall z \forall v \Big(\big(p\big(x, g\big(h(x,z)\big), z\big) \vee \neg q\big(i(x,z)\big)\big) \wedge r\Big(f\big(v, j(x,z)\big), j(x,z)\Big)\Big)$$

Allabschluß (betrifft jeweils **alle freien Variablen** – hier ist solche nur x)

Skolemisierung

Skolem-Normalform von F:

$$\forall x \forall z \forall v \Big(\big(p\big(x, g\big(h(x,z)\big), z\big) \vee \neg q\big(i(x,z)\big)\big) \wedge r\Big(f\big(v, j(x,z)\big), j(x,z)\Big)\Big)$$

Skolemisierungsweg:

Aus $\forall x_1 \ldots \forall x_n \exists y\, Q_1 z_1 \ldots Q_m z_m\, M$	Q_i **ist** $\exists$ **oder** $\forall$
mache $\forall x_1 \ldots \forall x_n\, Q_1 z_1 \ldots Q_m z_m\, M[y/f(x_1 \ldots x_n)]$	$M[y/f(x_1 \ldots x_n)]$ bedeutet das Ergebnis der in M durchgeführten Ersetzung von y durch **eine neue Funktion** f. Die Argumenten dieser Funktion sind genau all die allquantifizierten Variablen (hier - $x_1 \ldots x_n$), in denen Bindungsbereich die Variable y lag. Wie man sieht, die Stelligkeit solcher Funktion ist gleich der **Anzahl** von allen allquantifizierten Variablen, die vor der zu ersetzenden existenzquantifizierten Variablen stehen: sollte keine allquantifizierten Variablen vor so einer zu ersetzenden Variablen stehen ***(die oben erwähnte Anzahl ist dann einfach*** **0** ***)***, wird diese zu ersetzende Variable durch eine neue 0-stellige Funktion (anders gesagt – **durch eine neue** Konstante) ersetzt.
wiederhole, bis alle $\exists$ eliminiert	