CS 540: INTRODUCTION TO ARTIFICIAL INTELLIGENCE

HOMEWORK ASSIGNMENT #8

QUESTION 1: RESOLUTION PROOF IN PROPOSITIONAL LOGIC

Given the knowledge base

$$\begin{aligned}
KB &\equiv p \Rightarrow (q \Rightarrow r) \\
&\equiv \neg p \lor (q \Rightarrow r) && \text{(Implication Elimination)} \\
&\equiv \underbrace{\neg p \lor (\neg q \lor r)}_{\text{CNF}} && \text{(Implication Elimination)}
\end{aligned}$$

and the query β,

$$\begin{aligned}
\beta &\equiv (p \Rightarrow q) \Rightarrow (p \Rightarrow r) \\
&\equiv \neg(p \Rightarrow q) \lor (p \Rightarrow r) && \text{(Implication Elimination)} \\
&\equiv \neg(\neg p \lor q) \lor (\neg p \lor r) && \text{(Implication Elimination)} \\
&\equiv (p \land \neg q) \lor (\neg p \lor r) && \text{(de Morgan)} \\
&\equiv (\underbrace{p \lor \neg p}_{\text{TRUE}} \lor r) \land (\neg q \lor \neg p \lor r) && \text{(Distributivity of } \land \text{ over } \lor) \\
&\equiv \underbrace{\neg p \lor (\neg q \lor r)}_{\text{CNF}}
\end{aligned}$$

Adding $\neg\beta$ to KB to prove by ~~res~~ resolution

$a: \neg p \lor (\neg q \lor r) \rightarrow KB$

$b: \neg(\neg p \lor (\neg q \lor r)) \rightarrow \neg\beta$

$\Rightarrow a \land b = \text{FALSE} \quad \therefore KB \land \neg\beta \vdash \text{empty}$

Thus it is proved by contradiction that β logically follows the knowledge base (KB). KB entails β.

QUESTION 2: TRANSLATION TO FIRST ORDER LOGIC

1. i) Anything that jumps higher than a building is not a building, because buildings don't jump.

ii) FOL Variables - x, y ; Domain - Universal set of objects

iii) FOL Predicates -

- Building (x) : Returns TRUE if object x is a building
  Returns FALSE otherwise

- Jumps (x,y) : Returns TRUE if object x jumps higher than object y
  Returns FALSE otherwise
  If $x = y$, returns FALSE since x doesn't (cannot) jump higher than x.

iv) FOL Sentences:

$$\forall x \, \forall y \, [\text{Jumps}(x,y) \land \text{Building}(y)] \Rightarrow \neg \text{Building}(x)$$

↑
UNIVERSAL QUANTIFIER

2. i) Each of the 100 politicians is either honest or a liar. Atleast one of 100 politicians is honest. Of any two of the 100 politicians, atleast one of them is a liar.

ii) FOL Variables - x, y ; Domain - Universal set of objects (but are actually filtered by PREDICATE functions) (Politician(x))

iii) FOL Predicates -

Politician (x) - { Returns TRUE if x is one of the 100 politicians in the party
Returns FALSE otherwise }

Honest (x) - { Returns TRUE if ~~x ∈ set of 100 politicians and~~ x is honest.
Returns FALSE otherwise }

Liar (x) - { Returns TRUE if ~~x ∈ set of 100 politicians and~~ x is a liar
Returns FALSE otherwise }

iv) FOL Sentences:

→ $\forall x \; Politician(x) \Rightarrow [Honest(x) \vee Liar(x)]$

→ $\exists x \; Politician(x) \wedge Honest(x)$

→ $\forall x,y \; [Politician(x) \wedge Politician(y) \wedge (x != y)] \Rightarrow [Liar(x) \vee Liar(y)]$

QUESTION 3: HIERARCHICAL CLUSTERING

1)

	MADISON	SEATTLE	BOSTON	VANCOUVER	WINNIPEG	MONTREAL
MADISON	0	1617	931	1654	597	800
SEATTLE	1617	0	2486	121	1153	2283
BOSTON	931	2486	0	2501	1344	250
VANCOUVER	1654	121	2501	0	1159	2291
WINNIPEG	597	1153	1344	1159	0	1132
MONTREAL	800	2283	250	2291	1132	0

NOTE: Distance (fly-distance) between the cities are mentioned in miles (rounded to the nearest mile)

2) Let us denote the six cities with the following shorthand notations.

Ma: MADISON

S: SEATTLE

B: BOSTON

V: VANCOUVER

W: WINNIPEG

Mo: MONTREAL

ITERATION-1 :

CLUSTERS -

Ma S B V W Mo

i) Closest pair of clusters : (SEATTLE) & (VANCOUVER) S & V

ii) Distance (based on complete linkage) : 121 miles

iii) Clusters at the end of the iteration

Ma S,V B W Mo

ITERATION -2 :

CLUSTERS -

Ma S,V B W Mo

i) Closest pair of clusters : (BOSTON) & (MONTREAL) B & Mo

ii) Distance (based on complete linkage) : 250 miles

iii) Clusters at the end of the iteration

Ma S,V B,Mo W

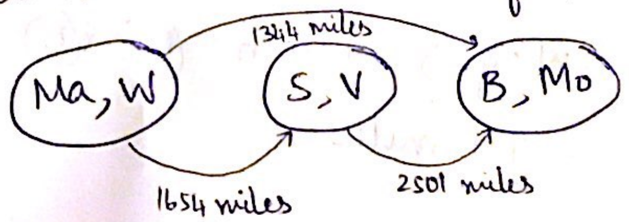
ITERATION -3 :

Ma S,V B,Mo W

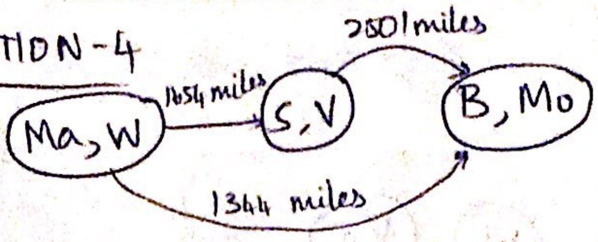
i) Closest pair of clusters : (MADISON) & (WINNIPEG) Ma & W

ii) Distance (based on complete linkage) : 597 miles

iii) Clusters at the end of the iteration



ITERATION-4



i) Closest pair of clusters: (MADISON, WINNIPEG) & (BOSTON, MONTREAL)

Ma, W & B, Mo

ii) Distance (based on complete linkage): 1344 miles.

iii) Clusters at the end of the iteration:

Ma, W, B, Mo   S, V

Clustering terminates as two clusters are produced.

---

3) At no point, a US city and a canadian city can be put in the SAME cluster

ITERATION-1

CLUSTERS

Ma   S   B   V   W   Mo

i) Closest pair of clusters: (MADISON) & (BOSTON)   Ma & B

$S \xrightarrow{121 \text{ miles}} V$

$B \xrightarrow{251 \text{ miles}} Mo$

$Ma \xrightarrow{597 \text{ miles}} W$

$Ma \xrightarrow{800 \text{ miles}} Mo$

→ The pairs of clusters are considered infinitely apart as the complete linkage is due to cities in two different countries

ii) Distance (based on complete linkage) between Ma & B

: 931 miles

iii) Clusters at the end of the iteration:

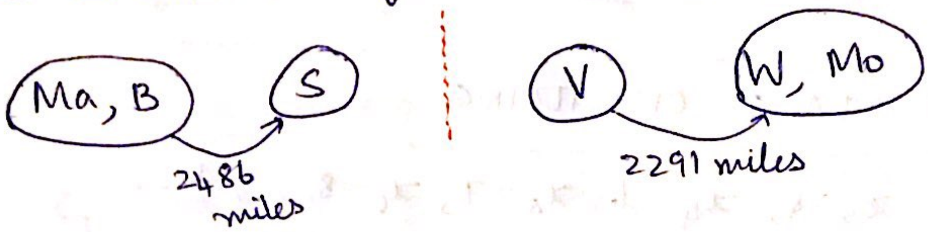
ITERATION - 2

CLUSTERS

i) Closest pair of clusters: (WINNIPEG) & (MONTREAL) : W & Mo

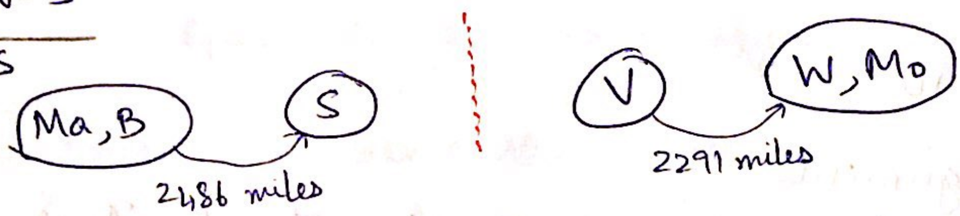
ii) Distance (based on complete linkage) : 1132 miles

iii) Clusters at the end of the iteration



ITERATION - 3

CLUSTERS

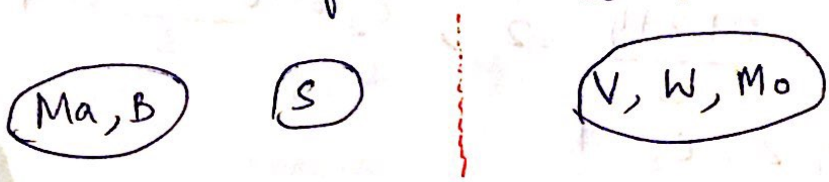


i) Closest pair of clusters : (VANCOUVER) & (WINNIPEG, MONTREAL)

V & W, Mo

ii) Distance (based on complete linkage) : 2291 miles

iii) Clusters at the end of the iteration:



ITERATION - 4:

CLUSTERS

(Ma, B) $\xrightarrow{2486 \text{ miles}}$ (S) | (V, W, Mo)

i) Closest pair of clusters: (MADISON, BOSTON) & (SEATTLE)

(Ma, B) & (S)

ii) Distance (based on complete linkage): 2486 miles

iii) Clusters at the end of the iteration

(Ma, B, S) | (V, W, Mo)

Clustering terminates as two clusters are produced.

QUESTION 4: K-MEANS CLUSTERING

$x_1 = 0$, $x_2 = 2$, $x_3 = 4$, $x_4 = 6$, $x_5 = 7$, $x_6 = 8$; $k = 2$

i)

ITERATION - 1:

$c_1 = 1$; $c_2 = 10$

1) Cluster assignments

$x_1 = 0$ $\quad x_2 = 2$ $\quad x_3 = 4$ $\quad x_4 = 6$ $\quad x_5 = 7$ $\quad x_6 = 8$

$y_1 = 1$ $\quad y_2 = 1$ $\quad y_3 = 1$ $\quad y_4 = 2$ $\quad y_5 = 2$ $\quad y_6 = 2$

2) Updated cluster centers

$$c_1 = \frac{0+2+4}{3} = 2 \; ; \; c_2 = \frac{6+7+8}{3} = 7$$

$c_1 = 2$; $c_2 = 7$

3) Energy at the end of the iteration

$$\text{Energy} = \sum \| c_i - x_{y_i} \|^2$$

$$\text{Energy} = (2-0)^2 + (2-2)^2 + (2-4)^2 + (7-6)^2 + (7-7)^2 + (7-8)^2$$

$$= 4 + 0 + 4 + 1 + 0 + 1 = 10.$$

ITERATION - 2:

$c_1 = 2$; $c_2 = 7$

1) Cluster assignments

$x_1 = 0$ $x_2 = 2$ $x_3 = 4$ $x_4 = 6$ $x_5 = 7$ $x_6 = 8$

$y_1 = 1$ $y_2 = 1$ $y_3 = 1$ $y_4 = 2$ $y_5 = 2$ $y_6 = 2$

2) Updated cluster centers

$$c_1 = \frac{0+2+4}{3} = 2 \; ; \quad c_2 = \frac{6+7+8}{3} = 7$$

$c_1 = 2$; $c_2 = 7$ $\Rightarrow$ Energy $= 10$

Cluster centers and hence the cluster assignments no longee change. therefore, the clustering process terminates.

---

2) $c_1 = 1$; $c_2 = 2$

ITERATION - 1

1) Cluster assignments

$x_1 = 0$ $x_2 = 2$ $x_3 = 4$ $x_4 = 6$ $x_5 = 7$ $x_6 = 8$

$y_1 = 1$ $y_2 = 2$ $y_3 = 2$ $y_4 = 2$ $y_5 = 2$ $y_6 = 2$

2) Updated cluster centers

$$c_1 = \frac{0}{1} = 0 \; ; \quad c_2 = \frac{2+4+6+7+8}{5} = 5.4$$

3) Energy at the end of the iteration $(c_1 = 0;\ c_2 = 5.4)$

$$\text{Energy} = \sum \| c_i - x_{y_i} \|^2$$

$$= (0-0)^2 + (5.4-2)^2 + (5.4-4)^2 + (5.4-6)^2 + (5.4-7)^2 + (5.4-8)^2$$

$$= 3.4^2 + 1.4^2 + 0.6^2 + 1.6^2 + 2.6^2$$

$$= 23.2$$

ITERATION - 2: $(c_1 = 0,\ c_2 = 5.4)$

1) Cluster assignments

$x_1 = 0 \quad x_2 = 2 \quad x_3 = 4 \quad x_4 = 6 \quad x_5 = 7 \quad x_6 = 8$

$y_1 = 1 \quad y_2 = 1 \quad y_3 = 2 \quad y_4 = 2 \quad y_5 = 2 \quad y_6 = 2$

2) Updated cluster centers

$$c_1 = \frac{0+2}{2} = 1 \ ; \quad c_2 = \frac{4+6+7+8}{4} = 6.25$$

$$c_1 = 1;\ c_2 = 6.25$$

3) Energy at the end of the iteration

$$\text{Energy} = \sum \| c_i - x_{y_i} \|^2$$

$$= (1-0)^2 + (1-2)^2 + (6.25-4)^2 + (6.25-6)^2 + (6.25-7)^2 + (6.25-8)^2$$

$$= 1 + 1 + 2.25^2 + 0.25^2 + 0.75^2 + 1.75^2$$

$$= 10.75$$

ITERATION-3 $(c_1 = 1 ; c_2 = 6.25)$

1) Cluster assignments

$x_1 = 0$ $x_2 = 2$ $x_3 = 4$ $x_4 = 6$ $x_5 = 7$ $x_6 = 8$

$y_1 = 1$ $y_2 = 1$ $y_3 = 2$ $y_4 = 2$ $y_5 = 2$ $y_6 = 2$

2) Updated cluster centers

$$c_1 = \frac{0+2}{2} = 1 \; ; \quad c_2 = \frac{4+6+7+8}{4} = 6.25$$

$$c_1 = 1 ; \; c_2 = 6.25$$

cluster centers and hence the cluster assignments no longer change and energy also stays at 10.75. Thus, the clustering process terminates.

---

3) The k-means solution with $c_1 = 1, c_2 = 10$ is better, because

- it provides a lower energy than the clustering with $c_1 = 1, c_2 = 2$
- the aim of k-means clustering is to minimize the pairwise squared deviations of points within the same cluster (i.e. error minimization problem) which $c_1 = 1, c_2 = 10$ achieves better.