CS 540: Introduction to Artificial Intelligence
Homework Assignment # 8

Yien Xu

Question 1: Resolution Proof in Propositional Logic [25 points]

Given the knowledge base

$$p \implies (q \implies r)$$

use resolution to prove the query

$$(p \implies q) \implies (p \implies r).$$

Be sure to show what you convert to CNF and how (do not skip steps), and how you perform each resolution step.

1. Convert α to CNF

$$\begin{aligned} \alpha : A &= p \Rightarrow (q \Rightarrow r) \\ &\equiv p \Rightarrow (\neg q \vee r) \\ &\equiv \neg p \vee (\neg q \vee r) \\ &\equiv \neg p \vee \neg q \vee r \end{aligned}$$

2. Convert $\neg\beta$ to CNF

$$\begin{aligned} \beta : B &= \neg\left((p \Rightarrow q) \Rightarrow (p \Rightarrow r)\right) \\ &\equiv \neg((\neg p \vee q) \Rightarrow (\neg p \vee r)) \\ &\equiv \neg(\neg(\neg p \vee q) \vee (\neg p \vee r)) \\ &\equiv \neg((p \wedge \neg q) \vee (\neg p \vee r)) \\ &\equiv \neg(p \wedge \neg q) \wedge \neg(\neg p \vee r) \\ &\equiv (\neg p \vee q) \wedge (p \wedge \neg r) \\ &\equiv (\neg p \vee q) \wedge p \wedge \neg r \end{aligned}$$

3. Knowledge Base

   $1 : \neg p \vee \neg q \vee r$

   $2 : \neg p \vee q$

   $3 : p$

   $4 : \neg r$

4. Resolution

   $5 : 2, 3 = (\neg p \vee q) \wedge p \equiv (p \wedge \neg p) \vee (p \wedge q) \equiv p \wedge q$

   $6 : q$ (from 5)

$7 : 1, 3, 4 = (\neg p \vee \neg q \vee r) \wedge p \wedge \neg r \equiv ((\neg p \wedge p) \vee (\neg q \wedge p) \vee (r \wedge p)) \wedge \neg r \equiv ((\neg q \wedge p) \vee (r \wedge p)) \wedge \neg r \equiv (\neg q \wedge p \wedge \neg r) \vee (r \wedge \neg r \wedge p) \equiv \neg q \wedge p \wedge \neg r$

$8 : \neg q$ (from 7)

null : $6, 8 = q \wedge \neg q$

Therefore, $(p \Rightarrow (q \Rightarrow r)) \mid = ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$

Question 2: Translation to First Order Logic [25 points]

Here are two riddles.

1. Question: What jumps higher than a building? Answer: Everything, buildings don't jump.
2. There is a party of 100 politicians. All of them are either honest or liars. You know two things: 1. At least one of them is honest. 2. If you take any two politicians, at least one of them is a liar.

For each riddle separately, do the following:

1. Write a plain English explanation for the riddle. This explanation should correspond to your logic statement later. For example, the first riddle can probably be stated as "Everything that is not a building jumps higher than a building," or "Everything that can jump jumps higher than a building." But the explanation "Everything jumps higher than a building" might be problematic if everything includes that building itself. Use your judgment to create a concise but correct statement. Remove non-essential details. If you think the riddle is ambiguous, explain why so and which interpretation you picked.
2. Define your First Order Logic (FOL) variables and their domains.
3. Define your FOL predicates and functions. Make sure you specify their values for ALL their input combinations.
4. Give the FOL sentences.

Riddle 1

1. Everything that can jump and that is not a building can jump higher than a building.
2. Variable domain: all things in the world
3. building(x): x is a building; jump(x): x can jump; higher(x, y): x jumps **strictly** higher than y.
4. $(\forall x, y)$ jump$(x) \wedge \neg$building$(x) \wedge$ building$(y) \Rightarrow$ higher(x, y)

Riddle 2

1. Among a party of 100 politicians who are either honest or liar, there exists at least one honest politician, and for any two politicians, if one of them is honest, then the other is not honest.
2. Variable refers to a politician. The domain P is the party of 100 politicians
3. $H(x)$: x is honest.
4. $(\exists x \in P\ \forall m, n \in P)\ H(x) \wedge (H(m) \Rightarrow \neg H(n)) \wedge \neg(m = n)$

Question 3: Hierarchical Clustering [25 points]

Consider the following six major cities. In the US: Madison, Seattle, Boston; and in Canada: Vancouver, Winnipeg, Montreal.

1. Create a 6×6 table with the distances between the cities. Consult the website `https://www.distance-cities.com`. Use the pink "fly distance" (the shorter distance), not the blue distance by car. Use miles and round to the nearest mile.

	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

Distances between the cities

2. Use hierarchical clustering with complete linkage to produce TWO clusters by hand. Specifically, show the following in each iteration: (1) the closest pair of clusters; (2) the distance between them as defined by complete linkage; (3) all clusters at the end of that iteration.

   {Madison} {Seattle} {Boston} {Vancouver} {Winnipeg} {Montreal}

   (a) Dist(Seattle, Vancouver) = 121
   {Madison} {Seattle, Vancouver} {Boston} {Winnipeg} {Montreal}

   (b) Dist(Boston, Montreal) = 250
   {Madison} {Seattle, Vancouver} {Boston, Montreal} {Winnipeg}

   (c) Dist(Winnipeg, Madison) = 597
   {Madison, Winnipeg} {Seattle, Vancouver} {Boston, Montreal}

   (d) Dist({Madison, Winnipeg}, {Boston, Montreal}) = 1344
   {Madison, Winnipeg, Boston, Montreal} {Seattle, Vancouver}

3. Now repeat the above question, but with the following constraint: at no point should a US city and a Canadian city be put in the same cluster. Equivalently, whenever the complete linkage between two clusters is due to two cities in different countries, treat the two clusters as infinity apart, regardless of what other cities are in those two clusters. You still need to show all steps.

   {Madison} {Seattle} {Boston} {Vancouver} {Winnipeg} {Montreal}

   (a) Dist(Boston, Madison) = 931
   {Madison, Boston} {Seattle} {Vancouver} {Winnipeg} {Montreal}

   (b) Dist(Winnipeg, Montreal) = 1132
   {Madison, Boston} {Seattle} {Vancouver} {Winnipeg, Montreal}

   (c) Dist(Vancouver, {Winnipeg, Montreal}) = 2291
   {Madison, Boston} {Seattle} {Vancouver, Winnipeg, Montreal}

   (d) Dist(Seattle, {Madison, Boston}) = 2486
   {Madison, Boston, Seattle} {Vancouver, Winnipeg, Montreal}

Question 4: K-means Clustering [25 points]

Given the following six items in 1D: $x_1 = 0, x_2 = 2, x_3 = 4, x_4 = 6, x_5 = 7, x_6 = 8$, perform k-means clustering to obtain $k = 2$ clusters by hand. Specifically,

1. Start from initial cluster centers $c_1 = 1, c_2 = 10$. Show your steps for all iterations: (1) the cluster assignments $y_1, \ldots, y_6$; (2) the updated cluster centers at the end of that iteration; (3) the energy at the end of that iteration.

   $c_1 = 1, c_2 = 10$

   (a) $y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 2, y_5 = 2, y_6 = 2$
   $c_1 = \frac{0+2+4}{3} = 2, c_2 = \frac{6+7+8}{3} = 7$
   $E = |0-2|^2 + |2-2|^2 + |4-2|^2 + |6-7|^2 + |7-7|^2 + |8-7|^2 = 4+4+1+1 = 10$

   (b) $y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 2, y_5 = 2, y_6 = 2$
   $c_1 = \frac{0+2+4}{3} = 2, c_2 = \frac{6+7+8}{3} = 7$
   $E = |0-2|^2 + |2-2|^2 + |4-2|^2 + |6-7|^2 + |7-7|^2 + |8-7|^2 = 4+4+1+1 = 10$
   We should stop.

2. Repeat the above but start from initial cluster centers $c_1 = 1, c_2 = 2$.

   $c_1 = 1, c_2 = 2$

   (a) $y_1 = 1, y_2 = 2, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
   $c_1 = \frac{0}{1} = 0, c_2 = \frac{2+4+6+7+8}{5} = 5.4$
   $E = |0-0|^2 + |2-5.4|^2 + |4-5.4|^2 + |6-5.4|^2 + |7-5.4|^2 + |8-5.4|^2 = 23.2$

   (b) $y_1 = 1, y_2 = 1, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
   $c_1 = \frac{0+2}{2} = 1, c_2 = \frac{4+6+7+8}{4} = 6.25$
   $E = |0-1|^2 + |2-1|^2 + |4-6.25|^2 + |6-6.25|^2 + |7-6.25|^2 + |8-6.25|^2 = 10.75$

(c) $y_1 = 1, y_2 = 1, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
$c_1 = \frac{0+2}{2} = 1, c_2 = \frac{4+6+7+8}{4} = 6.25$
$E = |0-1|^2 + |2-1|^2 + |4-6.25|^2 + |6-6.25|^2 + |7-6.25|^2 + |8-6.25|^2 = 10.75$
We should stop.

3. Which k-means solution is better? Why?

The first k-means solution is better because the energy is lower.