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.

Solution

Given KB and β

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

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

Step 1: Add $\neg\beta$ to KB

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

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

Step 2: Convert KB to CNF

$$p \implies (q \implies r) \quad \text{starting sentence}$$

$$p \implies (\neg q \vee r) \quad \text{implication elimination}$$

$$\neg p \vee (\neg q \vee r) \quad \text{implication elimination}$$

$$\neg p \vee \neg q \vee r \quad \text{associativity of } \vee$$

$$\neg((p \implies q) \implies (p \implies r)) \quad \text{starting sentence}$$

$$\neg((\neg p \vee q) \implies (\neg p \vee r)) \quad \text{implication elimination}$$

$$\neg(\neg(\neg p \vee q) \vee (\neg p \vee r)) \quad \text{implication elimination}$$

$$\neg((p \wedge \neg q) \vee (\neg p \vee r)) \quad \text{de Morgan}$$

$$\neg(p \wedge \neg q) \wedge \neg(\neg p \vee r) \quad \text{de Morgan}$$

$$(\neg p \vee q) \wedge (p \wedge \neg r) \quad \text{de Morgan}$$

$$(\neg p \vee q) \wedge p \wedge \neg r \quad \text{associativity of } \wedge$$

KB is

$$\begin{aligned} a :& \neg p \vee \neg q \vee r \\ b1 :& \neg p \vee q \\ b2 :& p \\ b3 :& \neg r \end{aligned}$$

Step 3: Resolve a and b1: $\neg p \vee r$
Step 4: Resolve above and b2: r
Step 5: Resolve above and b3: empty

Alternate Solution
Given KB and β
KB: $p \implies (q \implies r)$
$\beta :$ $(p \implies q) \implies (p \implies r)$

Step 1: Converting KB to CNF:

$$p \implies (\neg q \vee r) \quad \text{implication elimination}$$

$$\neg p \vee (\neg q \vee r) \quad \text{implication elimination}$$

$$\neg p \vee \neg q \vee r \quad \text{associativity of } \vee$$

Step 2: Converting β to CNF:

$$(\neg p \vee q) \implies (\neg p \vee r) \quad \text{implication elimination}$$

$$\neg(\neg p \vee q) \vee (\neg p \vee r) \quad \text{implication elimination}$$

$$(p \wedge \neg q) \vee (\neg p \vee r) \quad \text{de Morgan}$$

$$(p \vee \neg p \vee r) \wedge (\neg q \vee \neg p \vee r) \quad \text{distributivity of } \vee \text{ over } \wedge$$

As $(p \vee \neg p)$ is always true, the above expression reduces to

$$\neg q \vee \neg p \vee r$$

$$\neg p \vee \neg q \vee r \quad \text{Associativity of } \vee$$

Step 3: AS KB and β have the same CNF form, resolution of KB$\wedge\neg\beta$ will be an empty set

Grading:
5 points for attempting.
5 points if correctly mentioned most of the resolution steps.
5 points if no step is skipped.
10 points if proof is correct.

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.

Solution
The first riddle.
1. Everything that can jump, jumps higher than a building
2. variables: x is an object, y is an object
   domains: objects

---

3. building(x): is true if x is a building, else false
   jump(x): is true if x can jump, else false
   higher_than(x,y): is true if x can jump higher than y, else false
4. $\forall x \forall y \quad jump(x) \wedge building(y) \implies higher\_than(x, y)$

The second riddle.
1. There is at least one from a party of 100 politicians who is honest. If you randomly pick any two different politicians from the party, it's impossible that two of them are honest.
2. variables: x is a politician in the party, y is politician in the party
   domains: a party of 100 politicians
3. honest(x): is true if x is honest, else false
4. $\exists x \quad honest(x)$
   $\exists x \forall y \quad \neg(x = y) \implies \neg(honest(x) \wedge honest(y))$

Grading for each Riddle:
1.5 points for part 1 for logically correct description.
2 points for part 2.
2 points for describing predicates and functions.
2 points for specifying the value for all their input combinations.
5 points for part 4 (deduct two points if made some minor mistakes)

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.
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.
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.

Solution

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

*unit in miles

2.
Iteration 1:
the closest pair: {Seattle}, {Vancouver}
distance: 121 miles
clusters at the end of iteration 1: {Seattle, Vancouver}, {Madison}, {Boston}, {Winnipeg}, {Montreal}

Iteration 2:
the closest pair: {Boston}, {Montreal}
distance: 250 miles
clusters at the end of iteration 2: {Seattle, Vancouver}, {Boston, Montreal}, {Madison}, {Winnipeg}

Iteration 3:
the closest pair: {Madison}, {Winnipeg}
distance: 597 miles
clusters at the end of iteration 3: {Seattle, Vancouver}, {Boston, Montreal}, {Madison, Winnipeg}

Iteration 4:
the closest pair: {Boston, Montreal}, {Madison, Winnipeg}
distance: 1344 miles
clusters at the end of iteration 4: {Seattle, Vancouver}, {Boston, Montreal, Madison, Winnipeg}

3.
Iteration 1:
the closest pair: {Madison}, {Boston}
distance: 931 miles
clusters at the end of iteration 1: {Madison, Boston}, {Seattle}, {Vancouver}, {Winnipeg}, {Montreal}

Iteration 2:
the closest pair: {Winnipeg}, {Montreal}
distance: 1132 miles
clusters at the end of iteration 2: {Madison, Boston}, {Winnipeg, Montreal}, {Seattle}, {Vancouver}

Iteration 3:
the closest pair: {Winnipeg, Montreal}, {Vancouver}

distance: 2291 miles
clusters at the end of iteration 3: {Madison, Boston}, {Winnipeg, Montreal, Vancouver}, {Seattle}

Iteration 4:
the closest pair: {Madison, Boston}, {Seattle}
distance: 2486 miles
clusters at the end of iteration 4: {Madison, Boston, Seattle}, {Winnipeg, Montreal, Vancouver}

Grading:
9 points for part 1 (0.25 points for each entry of the table).
8 points for part 2, 2 points per iteration (0.5 point for closest pair, 0.5 point for distance and 1 point for cluster) .
8 points for part 3, 2 points per iteration (0.5 point for closest pair, 0.5 point for distance and 1 point for cluster) .

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.
2. Repeat the above but start from initial cluster centers $c_1 = 1, c_2 = 2$.
3. Which k-means solution is better? Why?

Solution
1.
Iteration 1:
cluster asignments: $y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 2, y_5 = 2, y_6 = 2$
updated centers: $c_1 = 2, c_2 = 7$
energy: 10

Iteration 2:
cluster asignments: $y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 2, y_5 = 2, y_6 = 2$
updated centers: $c_1 = 2, c_2 = 7$ (convergence)
energy: 10

2.

Iteration 1:
cluster asignments: $y_1 = 1, y_2 = 2, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
updated centers: $c_1 = 0, c_2 = 5.4$
energy: 23.2

Iteration 2:
cluster asignments: $y_1 = 1, y_2 = 1, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
updated centers: $c_1 = 1, c_2 = 6.25$
energy: 10.75

Iteration 3:
cluster asignments: $y_1 = 1, y_2 = 1, y_3 = 2, y_4 = 2, y_5 = 2, y_6 = 2$
updated centers: $c_1 = 1, c_2 = 6.25$ (convergence)
energy: 10.75

3. Logically, k-means attempts to minimize distortion (cost), so the first solution is better because it results in lower distortion

Grading:
8 points for part 1, 4 points for each iteration (1 point for cluster assignment, 2 points for centers and 1 point for energy).
12 points for part 2, 4 points for each iteration (1 point for cluster assignment, 2 points for centers and 1 point for energy).
5 points for part 3: (1)2 points for picking first solution.(2) 3 points for correct reason, (2 points if fewer iterations is mentioned as the reason) .