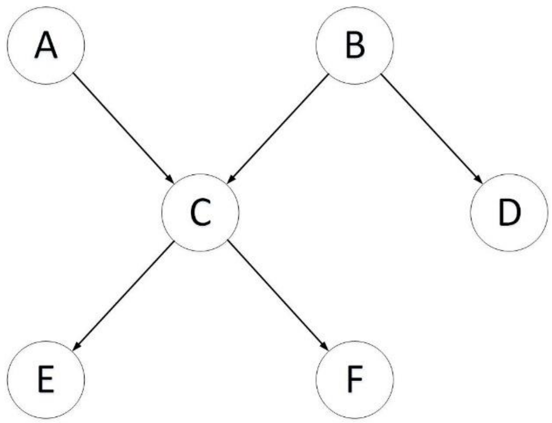
3346 Assignment 3

Zaid Albirawi

1. Consider the following belief network:



with Boolean variables (we write $A = true$ as a and $A = false$ as $\neg a$) and the following conditional probabilities:

$P(a) = 0.9$
$P(b) = 0.2$
$P(c \vee a, b) = 0.1$
$P(c \vee a, \neg b) = 0.8$
$P(c \vee \neg a, b) = 0.7$
$P(c \vee \neg a, \neg b) = 0.4$

$P(d \vee b) = 0.1$
$P(d \vee \neg b) = 0.8$
$P(e \vee c) = 0.7$
$P(e \vee \neg c) = 0.2$
$P(f \vee c) = 0.2$
$P(f \vee \neg c) = 0.9$

a. Compute $P(e)$ using VE. You should first prune irrelevant variables. Show the factors that are created from a given elimination ordering.

A	B	C	Probability
T	T	T	0.1
T	T	F	0.9
T	F	T	0.8
T	F	F	0.2
F	T	T	0.7
F	T	F	0.3
F	F	T	0.4
F	F	F	0.6

Firstly, eliminate A. This operation is performed by multiplying the probability of A by the above table and summing the alike B and C rows.

B	C	Probability
T	T	0.9(0.1) + 0.1(0.7) = 0.16
T	F	0.9(0.9) + 0.1(0.3) = 0.84
F	T	0.9(0.8) + 0.1(0.4) = 0.76
F	F	0.9(0.2) + 0.1(0.6) = 0.24

Secondly, eliminate B. This operation is performed by multiplying the probability of B by the above table and summing the alike C rows.

C	Probability
T	0.2(0.16) + 0.8(0.76) = 0.64
F	0.2(0.84) + 0.8(0.24) = 0.36

Furthermore, eliminate the variable C from the following table to find the probability of E.

C	E	Probability
T	T	0.7
T	F	0.3
F	T	0.2
F	F	0.8

Eliminate C by multiplying the probability of C by the above table and summing the alike E rows.

E	Probability
T	0.64(0.7) + 0.36(0.2) = 0.52
F	0.64(0.3) + 0.36(0.8) = 0.48

Therefore, the probability of e is 0.52

b. Suppose you want to compute $P(e \vee \neg f)$ using VE. How much of the previous computation can be reused? Show the factors that are different from those in part (a).

Firstly, start by multiplying the $P(e \vee c)$ and $P(f \vee c)$ tables.

C	E	Probability
T	T	0.7
T	F	0.3
F	T	0.2
F	F	0.8

C	F	Probability
T	T	0.2
T	F	0.8
F	T	0.9
F	F	0.1

C	E	F	Probability
T	T	T	0.7 * 0.2 = 0.14
T	T	F	0.7 * 0.8 = 0.56
T	F	T	0.3 *0.2 = 0.06
T	F	F	0.3 * 0.8 = 0.24
F	T	T	0.2 * 0.9 = 0.18
F	T	F	0.2 * 0.1 = 0.02
F	F	T	0.8 * 0.9 = 0.72
F	F	F	0.8 * 0.1 = 0.08

Multiple the table by the probability of C computed in question 1.a, and eliminate the variable C.

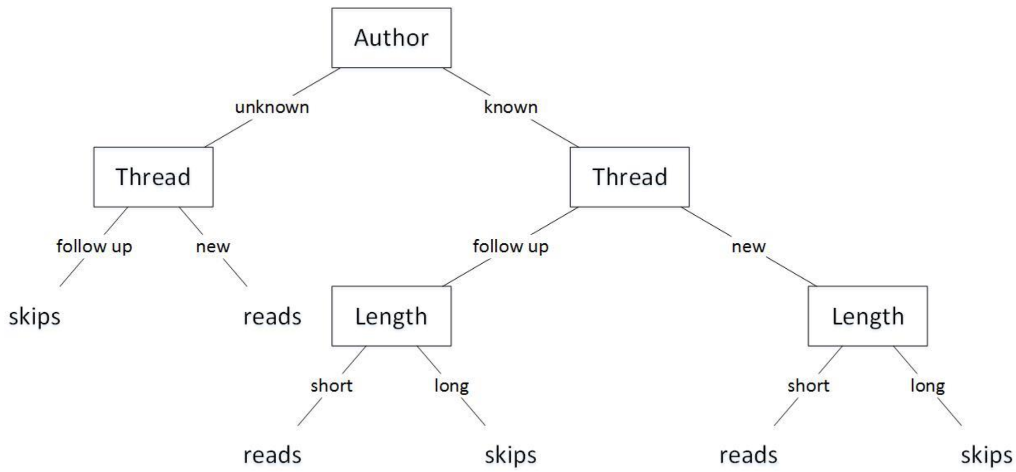
C	Probability
T	0.64
F	0.36

E	F	Probability
T	T	0.64(0.14) + 0.36(0.18) = 0.1544
T	F	0.64(0.56) + 0.36(0.02) = 0.3656
F	T	0.64(0.06) + 0.36(0.72) = 0.2976
F	F	0.64(0.24) + 0.36(0.08) = 0.1824

Therefore, the $P(e \vee \neg f) = 0.3656$.

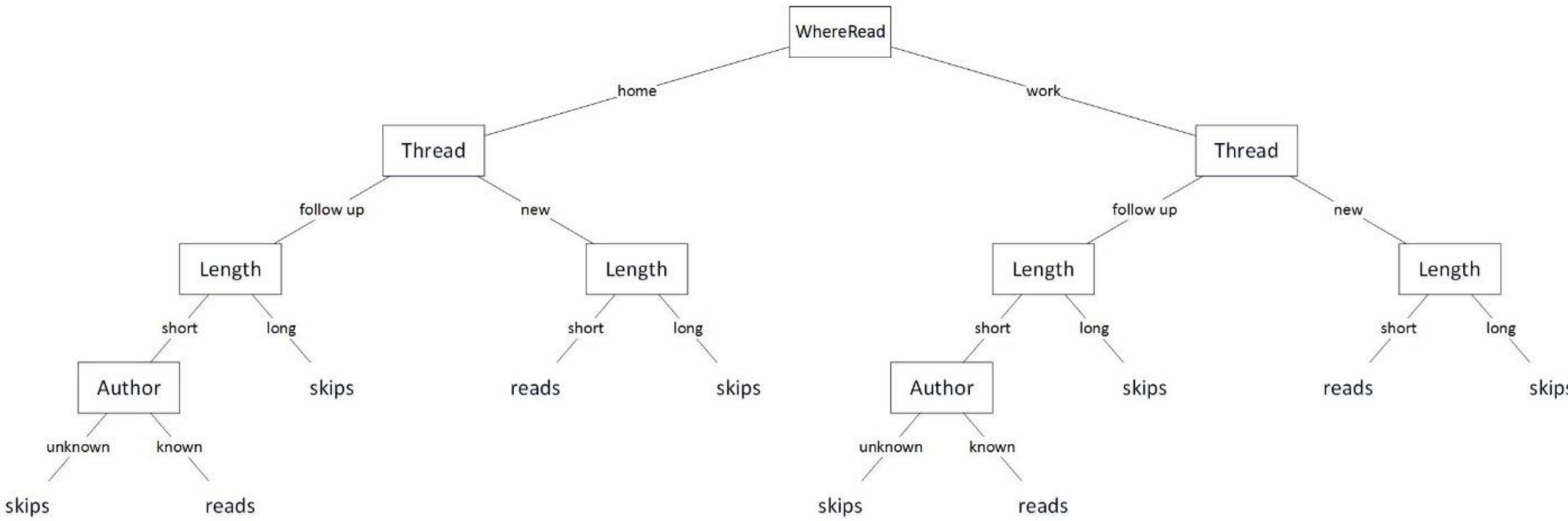
2. Consider the decision-tree learning algorithm of Figure 7.5 (page 300) and data of Figure 7.1 (page 289). Suppose, for this question, the stopping criterion is that all the examples have the classification. The tree of Figure 7.4 (page 298) was built by selecting a feature that gives the maximum information gain. This question considers what happens when a different feature is selected.

   a. Suppose you change the algorithm to always select the first element of the list of features. What tree is found when the features are in the order [*Author, Thread, Length, WhereRead*]? Does this tree represent a different function than that found with the maximum information gain split? Explain.



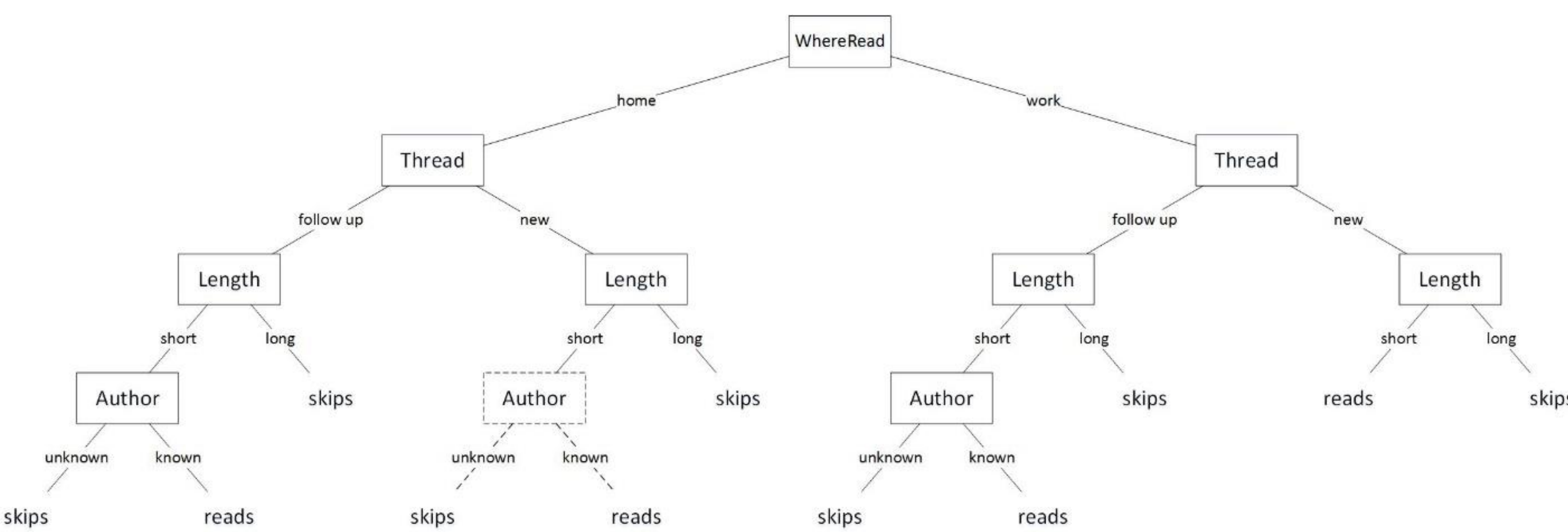
   The tree represents a different function from the one found with the maximum information gain split. The maximum information split tree results in skipping both e_{19} and e_{20}. While this tree results in skipping e_{20} and reading e_{19}. This is due to the fact that the maximum information gain split tree skips all long messages.

b. What tree is found the features are in the order [*WhereRead, Thread, Length, Author*]? Does this tree represent different function than that found with the maximum information gain split or the one given for the preceding part? Explain.

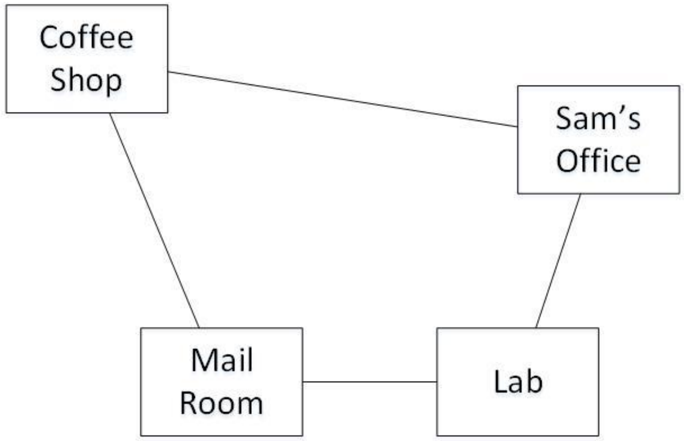


The tree represents a similar function to the one found with the maximum information gain split but a different function from the tree generated in the previous part of this question. This tree and the maximum information split tree result in skipping both e_{19} and e_{20}. While the tree generated in the previous question results in skipping e_{20} and reading e_{19}. This is because the tree generated in the previous question does not use the length attribute as one of its deciding factors when deciding on an entry with an unknown author and a follow up thread. This problem is caused by given the machine a small set of training examples that does not train it for all possible entries combinations.

c. Is there a tree that correctly classifies the training examples but represents a different function than those found by the preceding algorithms? If so, give it. If not, explain why.



3. Consider the planning domain:



Features to describe states

$RLoc$ — Rob's location
RHC — Rob has coffee
SWC — Sam wants coffee
MW — Mail is waiting
RHM — Rob has mail

Actions

mc — move clockwise
mcc — move counterclockwise
puc — pickup coffee
dc — deliver coffee
pum — pickup mail
dm — deliver mail

a. Give the feature-based representation of the *MW* and *RHM* features.

$$MW' = mw \leftarrow mw \wedge Act \neq pum$$
$$MW' = \overline{mw} \leftarrow mw \wedge Act = pum$$

$$RHM' = rhm \leftarrow mw \wedge Act = pum$$
$$RHM' = rhm \leftarrow rhm \wedge Act \neq dm$$
$$RHM' = \overline{rhm} \leftarrow rhm \wedge Act = dm$$

b. Give the STRIPS representations for the pick up mail and deliver mail actions.

pum
preconditions $[mr, mw]$
effects $[\overline{mw}, rhm]$

dm
preconditions $[off, rhm]$
effects $[\overline{rhm}]$

4. Suppose you have STRIPS representation for actions a_1 and a_2, and you want to define the STRIPS representation for the composite action $a_1; a_2$, which means that you do a_1 then do a_2.

   a. What is the effects list for this composite action?

      Let neg be an operator that provides the negation of the list elements, for examples $\text{effects}(x) = [e_1, e_2], neg\big(\text{effects}(x)\big) = [\overline{e_1}, \overline{e_2}]$.

      Let $effect(a_1, a_2) = (\text{effects}[a_2] \cup \text{effects}[a_1]) - (neg(\text{effects}[a_2]) \cap \text{effects}[a_1])$

      effects $\big[effect(a_1, a_2) - \big((\text{preconditions}[a_1; a_2]) \cap effect(a_1, a_2)\big)\big]$

   b. What are the preconditions for this composite action? You can assume that the preconditions are specified as a list $Variable = value$ pairs (rather than as arbitrary logical formulas).

      preconditions $\big[\text{preconditions}[a_1] \cup \big(\text{preconditions}[a_2] - (\text{preconditions}[a_2] \cap \text{effects}[a_1]\,)\big)\big]$

   c. Using the delivery robot domain of Example 8.1 (Page 350), give the STRIPS representation for the composite action $puc; mc$.

      preconditions $[cs, \overline{rhc}\,]$
      effects $[rhc, off]$

   d. Give STRIPS representation for the composite action $puc; mc; dc$ made up of three primitive actions.

      preconditions $[cs, \overline{rhc}\,]$
      effects $[off, \overline{swc}]$

   e. Give the STRIPS representation fir the composite action $mcc; puc; mc; dc$ made up of four primitive actions.

      preconditions $[off, \overline{rhc}\,]$
      effects $[\overline{swc}]$

5. Consider a domain with two individuals (✄ and ☎), two predicate symbols (p and q), and three constants (a, b and c). The knowledge base KB is defined by

$$p(X) \leftarrow q(X).$$
$$q(a).$$

a. Give one interpretation that is a model of KB.

$D = \{\text{✄}, \text{☎}\}$
$\varphi(a) = \text{☎}, \varphi(b) = \text{✄}, \varphi(c) = \text{✄}$
$\pi(p) = < \text{☎} > true, < \text{✄} > true$
$\pi(q) = < \text{☎} > true, < \text{✄} > true$

b. Give one interpretation that is not a model of KB.

$D = \{\text{✄}, \text{☎}\}$
$\varphi(a) = \text{☎}, \varphi(b) = \text{✄}, \varphi(c) = \text{✄}$
$\pi(p) = < \text{☎} > false, < \text{✄} > false$
$\pi(q) = < \text{☎} > false, < \text{✄} > false$

c. How many interpretations are there? Give a brief justification for your answer.

The number of ways two domains can be assigned to three constants is 2^3
The number ways two domains can be assigned to two predicates with a true and false values is 2^4

Therefore, $2^3 * 2^4 = 2^7 = 128$, there exists 128 different interpretations.

d. How many interpretations are models of KB? Give a brief justification for your answer.

Firstly, start by stating all possible values for $\{\text{✄}, \text{☎}\}$ in combination with the $p\ and\ q$ predicates.

$\pi(p) = < \text{☎} > true, < \text{✄} > true$
$\pi(p) = < \text{☎} > true, < \text{✄} > false$
$\pi(p) = < \text{☎} > false, < \text{✄} > true$
$\pi(p) = < \text{☎} > false, < \text{✄} > false$

$\pi(q) = < \text{☎} > true, < \text{✄} > true$
$\pi(q) = < \text{☎} > true, < \text{✄} > false$
$\pi(q) = < \text{☎} > false, < \text{✄} > true$
$\pi(q) = < \text{☎} > false, < \text{✄} > false$

Since $q(a)$ must be true for a model to exist, $\pi(q)$ must contain at least one true value at all times for us to able to obtain a KB model. Hence, this only works when we have $\pi(q) = < \text{☎} > true, < \text{✄} > true$ and $\pi(q) = < \text{☎} > true, < \text{✄} > false$ or $\pi(q) = < \text{☎} > false, < \text{✄} > true$, we can only have two because the last two contradict each other. Furthermore, for $p(X) \leftarrow q(X)$ to be true, we must choose instances of $\pi(p)$ that are compatible with the $\pi(q)$ instances that we choose earlier. Therefore, the number of interpretations that are models of KB is $\left(\frac{1}{4} * \frac{1}{2} + \frac{1}{4} * \frac{1}{4}\right) * 128 = 24$. The numbers correspond to $\left(\frac{1}{4}\ of\ \pi(q) * \frac{1}{2}\ of\ \pi(p) + \frac{1}{4}\ of\ \pi(q) * \frac{1}{4}\ of\ \pi(p)\right) * number\ of\ interpretations.$

6. Consider the following knowledge base:

$r(a)$.
$r(e)$.
$p(c)$.
$q(b)$.
$s(a,b)$.
$s(d,b)$.
$s(e,d)$.
$p(X) \leftarrow q(X) \wedge r(X)$.
$q(X) \leftarrow s(X,Y) \wedge q(Y)$.

Show the set of ground atomic consequences derivable from this knowledge base. Assume that a bottom-up proof procedure is used and that at each iteration the first applicable clause is selected in the order shown. Applicable constant substitutions are chosen in "alphabetic order" if more than one applies to a given clause; for example, if X/a and X/b are both applicable for a clause at some iteration, derive $q(a)$ first. In what order are consequences derived?

The set of all ground instances is

$r(a)$.
$r(e)$.
$p(c)$.
$q(b)$.
$s(a,b)$.
$s(d,b)$.
$s(e,d)$.
$p(a) \leftarrow q(a) \wedge r(a)$.
$p(b) \leftarrow q(b) \wedge r(b)$.
$p(c) \leftarrow q(c) \wedge r(c)$.
$p(d) \leftarrow q(d) \wedge r(d)$.
$p(e) \leftarrow q(e) \wedge r(e)$.
$q(a) \leftarrow s(a,a) \wedge q(a)$.
$q(a) \leftarrow s(a,b) \wedge q(b)$.
$q(a) \leftarrow s(a,c) \wedge q(c)$.
$q(a) \leftarrow s(a,d) \wedge q(d)$.
$q(a) \leftarrow s(a,e) \wedge q(e)$.
$q(b) \leftarrow s(b,a) \wedge q(a)$.
$q(b) \leftarrow s(b,b) \wedge q(b)$.
$q(b) \leftarrow s(b,c) \wedge q(c)$.
$q(b) \leftarrow s(b,d) \wedge q(d)$.
$q(b) \leftarrow s(b,e) \wedge q(e)$.
$q(c) \leftarrow s(c,a) \wedge q(a)$.
$q(c) \leftarrow s(c,b) \wedge q(b)$.
$q(c) \leftarrow s(c,c) \wedge q(c)$.
$q(c) \leftarrow s(c,d) \wedge q(d)$.
$q(c) \leftarrow s(c,e) \wedge q(e)$.

$q(d) \leftarrow s(d, a) \wedge q(a)$.
$q(d) \leftarrow s(d, b) \wedge q(b)$.
$q(d) \leftarrow s(d, c) \wedge q(c)$.
$q(d) \leftarrow s(d, d) \wedge q(d)$.
$q(d) \leftarrow s(d, e) \wedge q(e)$.
$q(e) \leftarrow s(e, a) \wedge q(a)$.
$q(e) \leftarrow s(e, b) \wedge q(b)$.
$q(e) \leftarrow s(e, c) \wedge q(c)$.
$q(e) \leftarrow s(e, d) \wedge q(d)$.
$q(e) \leftarrow s(e, e) \wedge q(e)$.

$\{\}$	initial state
$\{r(a)\}$	from: $r(a)$.
$\{r(a), r(e)\}$	from: $r(e)$.
$\{r(a), r(e), p(c)\}$	from: $p(c)$.
$\{r(a), r(e), p(c), q(b)\}$	from: $q(b)$.
$\{r(a), r(e), p(c), q(b), s(a, b)\}$	from: $s(a, b)$.
$\{r(a), r(e), p(c), q(b), s(a, b), s(d, b)\}$	from: $s(d, b)$.
$\{r(a), r(e), p(c), q(b), s(a, b), s(d, b), s(e, d)\}$	from: $s(e, d)$.

Let $\Sigma = \{r(a), r(e), p(c), q(b), s(a, b), s(d, b), s(e, d)\}$

$\{\Sigma, q(a)\}$	from: $q(a) \leftarrow s(a, b) \wedge q(b)$.
$\{\Sigma, q(a), p(a)\}$	from: $p(a) \leftarrow q(a) \wedge r(a)$.
$\{\Sigma, q(a), p(a), q(d)\}$	from: $q(d) \leftarrow s(d, b) \wedge q(b)$.
$\{\Sigma, q(a), p(a), q(d), q(e)\}$	from: $q(e) \leftarrow s(e, d) \wedge q(d)$.
$\{\Sigma, q(a), p(a), q(d), q(e), p(e)\}$	from: $p(e) \leftarrow q(e) \wedge r(e)$.

Therefore, the order the consequences are derived is,
$\{r(a), r(e), p(c), q(b), s(a, b), s(d, b), s(e, d), q(a), p(a), q(d), q(e), p(e)\}$.

7. In a manner similar to Example 12.21 (page 511), show derivations of the following queries:

   a. Ask $two\_doors\_east(r107, R)$.

      $yes(R) \leftarrow two\_doors\_east(r107, R)$
      resolve with $two\_doors\_east(E_1, W_1) \leftarrow imm\_east(E_1, M_1) \wedge imm\_east(M_1, W_1)$
      substitution: $\{E_1/r107, W_1/R\}$

      $yes(R) \leftarrow imm\_east(r107, M_1) \wedge imm\_east(M_1, R)$
      select leftmost conjunct
      resolve with $imm\_east(E_2, W_2) \leftarrow imm\_west(W_2, E_2)$
      substitution: $\{E_2/r107, W_2/M_1\}$

      $yes(R) \leftarrow imm\_west(M_1, r107) \wedge imm\_east(M_1, R)$
      select leftmost conjunct
      resolve with $imm\_west(r105, r107)$
      substitution: $\{M_1/r105\}$

      $yes(R) \leftarrow imm\_west(r105, r107) \wedge imm\_east(r105, R)$
      select leftmost conjunct
      resolve with $imm\_west(r105, r107)$
      substitution: $\{\ \}$

      $yes(R) \leftarrow imm\_east(r105, R)$
      resolve with $imm\_east(E_3, W_3) \leftarrow imm\_west(W_3, E_3)$
      substitution: $\{E_3/r105, W_3/R\}$

      $yes(R) \leftarrow imm\_west(R, r105)$
      resolve with $imm\_west(r103, r105)$
      substitution: $\{R/r103\}$

      $yes(r103) \leftarrow imm\_west(r103, r105)$
      resolve with $imm\_west(r103, r105)$
      substitution: $\{\ \}$

      $yes(r103) \leftarrow$

b. Ask $next\_door(R, r107)$.

$yes(R) \leftarrow next\_door(R, r107)$
resolve with $next\_door(E_1, W_1) \leftarrow imm\_east(E_1, W_1)$
substitution: $\{E_1/R, W_1/r107\}$

$yes(R) \leftarrow imm\_east(R, r107)$
resolve with $imm\_east(E_2, W_2) \leftarrow imm\_west(W_2, E_2)$
substitution: $\{W_2/r107, E_2/R\}$

$yes(R) \leftarrow imm\_west(r107, R)$
resolve with $imm\_west(r107, r109)$
substitution: $\{R/r109\}$

$yes(r109) \leftarrow imm\_west(r107, r109)$
resolve with $imm\_west(r107, r109)$
substitution: $\{\quad\}$

$yes(r109) \leftarrow$

$yes(R) \leftarrow next\_door(R, r107)$
resolve with $next\_door(W_3, E_3) \leftarrow imm\_west(W_3, E_3)$
substitution: $\{W_3/R, E_3/r107\}$

$yes(R) \leftarrow imm\_west(R, r107)$
resolve with $imm\_west(r105, r107)$
substitution: $\{R/r105\}$

$yes(r105) \leftarrow imm\_west(r105, r107)$
resolve with $imm\_west(r105, r107)$
substitution: $\{\quad\}$

$yes(r105) \leftarrow$

8. For each of the following pairs of atoms, either give a most general unifier or explain why one does not exist:
   a. $p(X, Y, a, b, W)$
      $p(E, c, F, G, F)$

      $E = \{X = E, Y = c, a = F, b = G, W = F\}$
      $S = \{ \quad \}$

      Select $X = E$

      $E = \{Y = c, a = F, b = G, W = F\}$
      $S = \{X/E\}$

      Select $Y = c$

      $E = \{a = F, b = G, W = F\}$
      $S = \{X/E, Y/c\}$

      Select $a = F$

      $E = \{b = G, W = a\}$
      $S = \{X/E, Y/c, a/F\}$

      Select $b = G$

      $E = \{W = a\}$
      $S = \{X/E, Y/c, a/F, b/G\}$

      Select $W = a$

      $E = \{ \quad \}$
      $S = \{X/E, Y/c, a/F, b/G, W/a\}$

      Therefore, the most general unifier is $\{X/E, Y/c, a/F, b/G, W/a\}$

b. $p(X, Y, Y)$
$p(E, E, F)$

$E = \{X = E, Y = E, Y = F\}$
$S = \{\quad\}$

Select $X = E$

$E = \{Y = E, Y = F\}$
$S = \{X/E\}$

Select $Y = E$

$E = \{E = F\}$
$S = \{X/E, Y/E\}$

Select $E = F$

$E = \{\quad\}$
$S = \{X/F, Y/F, E/F\}$

Therefore, the most general unifier is $\{X/F, Y/F, E/F\}$