8.11 Week 8 Homework Quiz

Kevin Offemaria (username: offemakp)

Attempt 2

Written: Mar 5, 2023 9:02 AM - Mar 5, 2023 9:04 AM

Submission View

Your quiz has been submitted successfully.

Question 1 **3 / 3 points**

Consider the following data set:

V1	V2	V3	V4	C(lass)
1	-1	-1	1	-1
1	1	1	-1	-1
1	-1	-1	-1	-1
1	1	1	1	-1
-1	-1	1	1	1
-1	1	-1	1	1
-1	1	-1	-1	1
-1	-1	1	-1	1

Which variable, Vi, would be selected by a decision tree learner of the type described in the text/lecture as the root of the decision tree?

- [ ] V1
- [ ] V2
- [ ] V3
- [ ] V4

Question 2 **3 / 3 points**

Consider the following data set:

V1	V2	V3	V4	C(lass)
1	-1	-1	1	-1
1	1	1	-1	-1
1	-1	-1	-1	-1
1	1	1	1	-1
-1	-1	1	1	1
-1	1	-1	1	1
-1	1	-1	-1	1
-1	-1	1	-1	1

Give the value of P(C=1|V2=1) as computed from the data table above (i.e., as a fraction or a floating point number; do not use pseudo-counts):

- [ ] 0.0
- [ ] 0.25
- [ ] 0.5
- [ ] 0.75
- [ ] 1.0
- [ ] 1.5

Question 3 **3 / 3 points**

Consider the following data set:

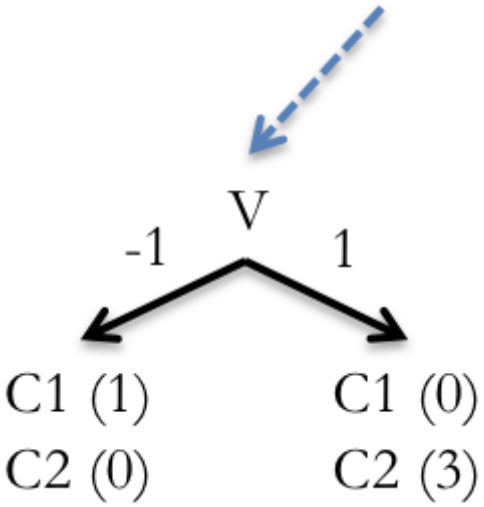
V1	V2	V3	V4	C(lass)
1	-1	-1	1	-1
1	1	1	-1	-1
1	-1	-1	-1	-1
1	1	1	1	-1
-1	-1	1	1	1
-1	1	-1	1	1
-1	1	-1	-1	1
-1	-1	1	-1	1

Give the value of P(C=1|V2=1, V3=1) as computed from the table above (as a fraction or a floating point number; do not use pseudo-counts):

- [ ] 0.0
- [ ] 0.25
- [ ] 0.5
- [ ] 0.75
- [ ] 1.0
- [ ] 1.5

Question 4 **3 / 3 points**

Assume that a decision tree has been constructed from training data, and it includes a node that tests on V at the frontier of the tree, with its left child yielding a prediction of class C1 (because the only training datum there is C1), and the right child predicting C2 (because the only training data there are C2). The situation is illustrated here:



Suppose that during subsequent use, it is found that

i) a large # of items (N > 1000) are classified to the node with the test on V

ii) 50% of these have V = -1 and 50% of these have V = 1

iii) post classification analysis shows that of the N items reaching the node during usage, 25% were C1 and 75% were C2

iv) of the 0.5 * N items that went to the left leaf during usage, 25% were C1 and 75% were C2

v) of the 0.5 * N items that went to the right leaf during usage, 25% were also C1 and 75% were C2

What is the error rate on the sample of N items that went to the sub-tree shown above?

- [ ] 0.0
- [ ] 0.25
- [ ] 0.5
- [ ] 0.75
- [ ] 1.0

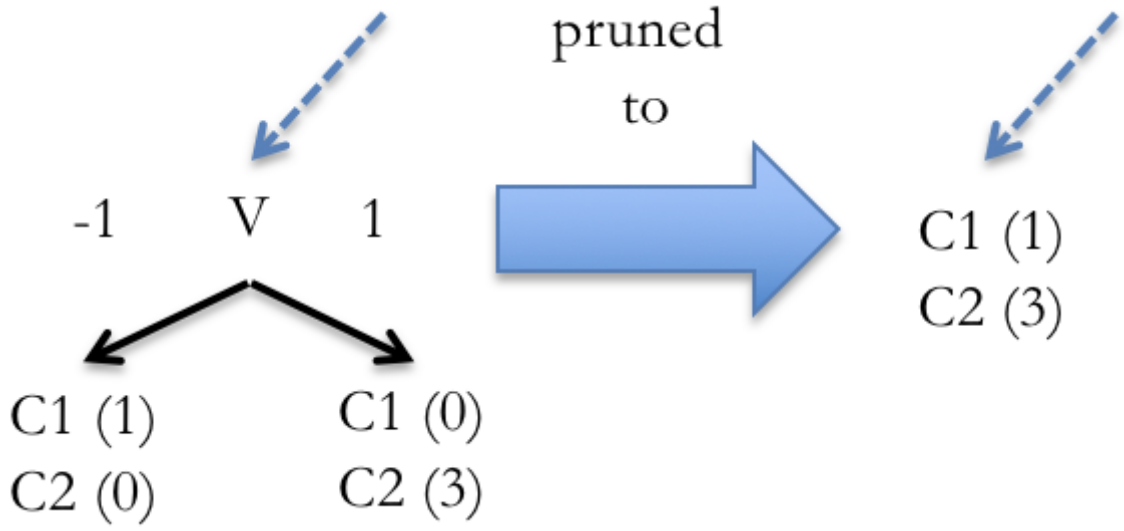
Question 5 **3 / 3 points**

What would the error rate on the same sample of N items have been if the sub-tree from the previous question (and reproduced here) had been pruned to not include the final test on V, but to rather be a leaf that predicted C2?

Suppose that during subsequent use, it is found that:

i) a large # of items (N > 1000) are classified to the leaf that predicts C2 (in the pruned case to the right below)

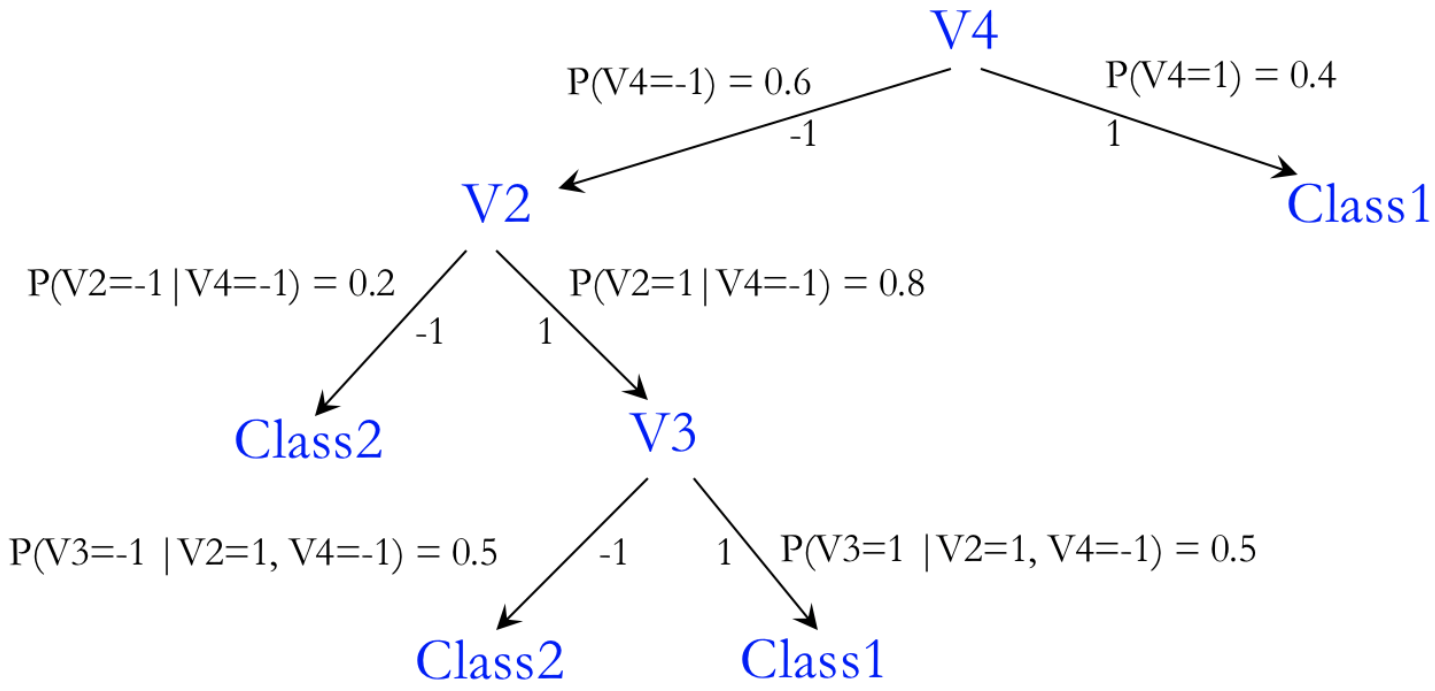
ii) post classification analysis shows that of the N items reaching the node during usage, 25% were C1 and 75% were C2



- [ ] 0.0
- [ ] 0.25
- [ ] 0.5
- [ ] 0.75
- [ ] 1.0

Question 6 **3 / 3 points**

Consider the following decision tree, where each variable (including the Class variable) is binary-valued. Each branch is labeled with the probability that the branch will be taken:

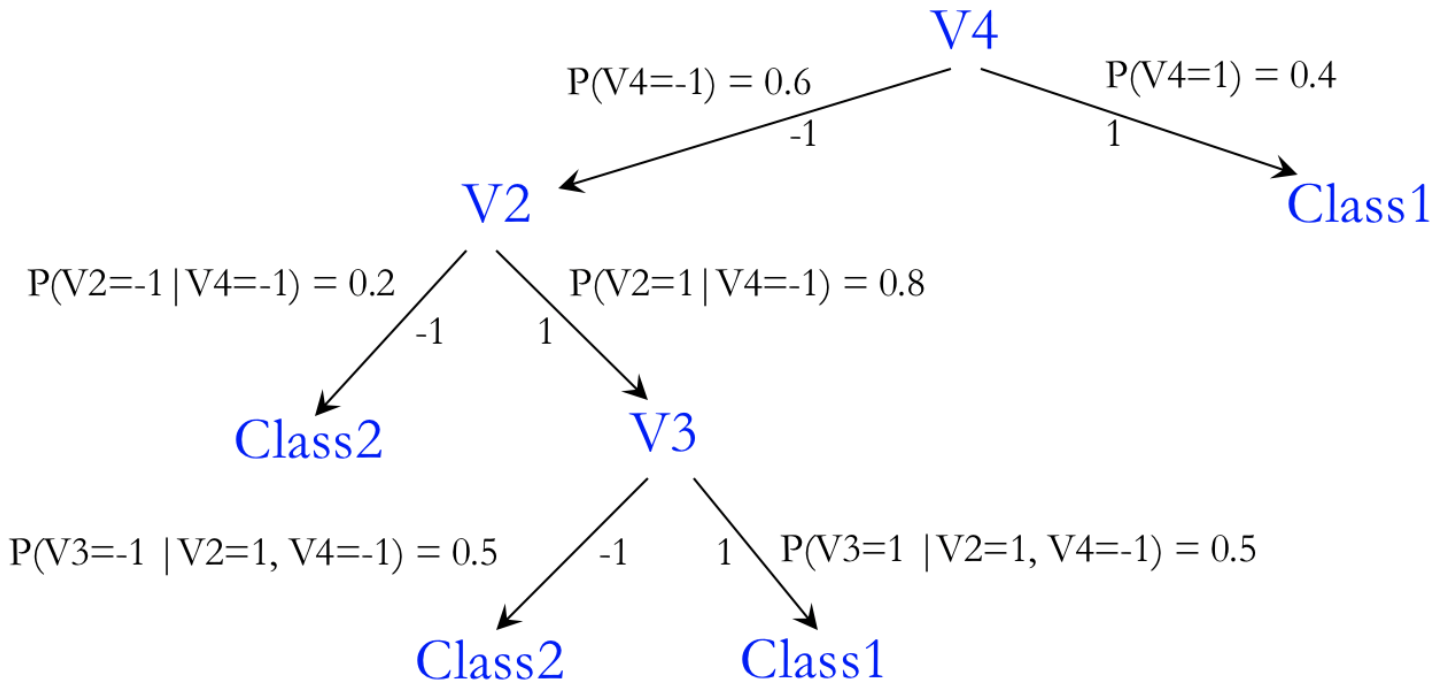


Give the expected number of internal nodes visited (i.e., the expected number of variable tests carried out) when classifying an arbitrary datum. (Internal nodes are nodes that are not leaves.)

- [ ] 1.0
- [ ] 1.62
- [ ] 2.08
- [ ] 3.0
- [ ] 3.28

Question 7 **2 / 2 points**

Consider the following decision tree, where each variable (including the Class variable) is binary-valued. Each branch is labeled with the probability that the branch will be taken:



When classifying a datum with all known values, the classification will proceed along exactly one path of the tree. For example, if the test datum was (V1= 1, V2 = -1, V3 = 1, V4 = -1), then classification would proceed down the leftmost path resulting in a Class2 prediction (with probability 1.0). But if there are missing values among the variable values of a test datum, classification may be nondeterministic (or probabilistic) -- classification may have to investigate multiple paths at nodes/variables where a test datum's value is not known.

Suppose that the test datum were (V1= 1, V2 = 1, V3 = ?, V4 = ?); that is, the values of V3 and V4 are not known. What would you conclude is the probability of a Class1 classification?

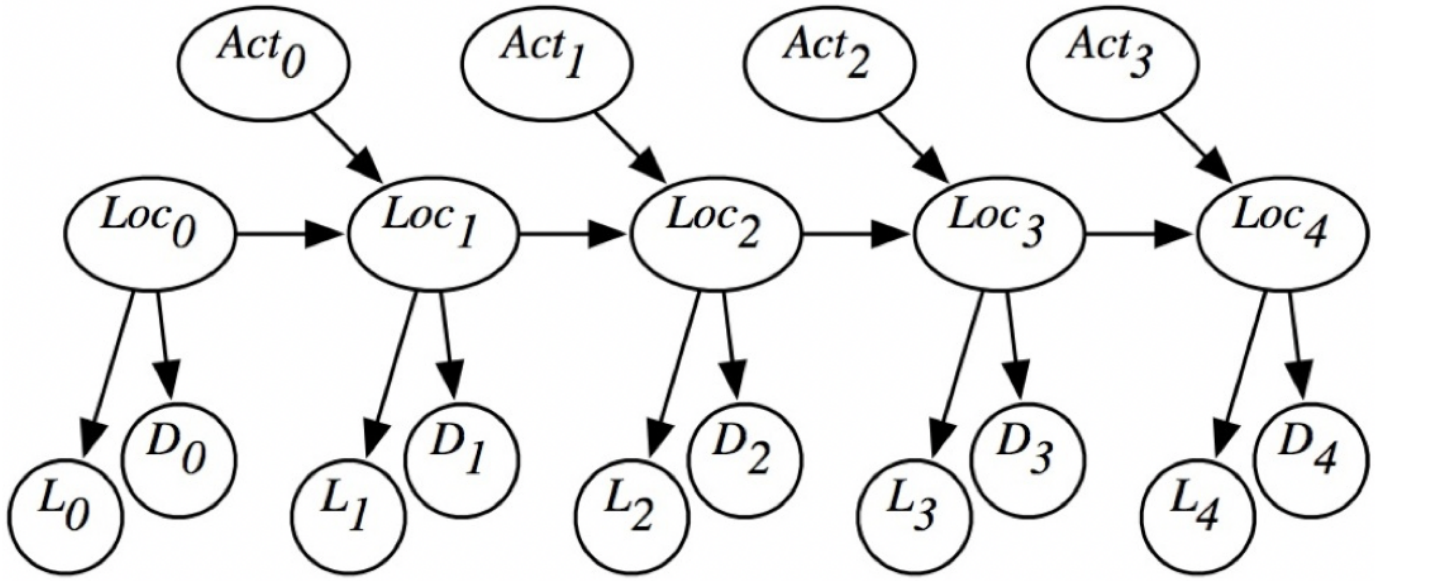
- [ ] 0.0
- [ ] 0.2
- [ ] 0.5
- [ ] 0.7
- [ ] 0.9
- [ ] 1.0

Question 8 **1 / 1 point**

Consider this Hidden Markov Model. Assume all variables are binary-valued. The domain of Loc_k is {c, ~c}, L_i is {l, ~l}, D_j is {d, ~d}, etc.

Make the assumption that both P(x|...) and P(~x|...) for each variable X are stored in probability tables.

Figure from Poole and Mackworth



Assume that the following are observed: L_0 = ~l, D_0 = d, Act_0 = a, L_1 = ~l, D_1 = ~d.

What two probabilities would have to be compared to determine which was more probable, Loc_1 = c or Loc_1 = ~c, given the observations above. **Choose two choices from the options below, where each selection should ONLY be in terms of probabilities found in the probability tables.**

- [ ] $P(Loc_1=c, L_0=\sim l, D_0=d, Act_0=a, L_1=\sim l, D_1=\sim d)$
- [ ] $P(Loc_1=\sim c, L_0=\sim l, D_0=d, Act_0=a, L_1=\sim l, D_1=\sim d)$
- [ ] $P(Loc_0=c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=c)\ P(D_0=d|Loc_0=c)\ P(Loc_1=c|Loc_0=c, Act_0=a)\ P(L_1=\sim l|Loc_1=c)\ P(D_1=\sim d|Loc_1=c)$

  +

  $P(Loc_0=\sim c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=\sim c)\ P(D_0=d|Loc_0=\sim c)\ P(Loc_1=c|Loc_0=\sim c, Act_0=a)\ P(L_1=\sim l|Loc_1=c)\ P(D_1=\sim d|Loc_1=c)$
- [ ] $P(Loc_0=c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=c)\ P(D_0=d|Loc_0=c)\ P(Loc_1=\sim c|Loc_0=c, Act_0=a)\ P(L_1=\sim l|Loc_1=\sim c)\ P(D_1=\sim d|Loc_1=\sim c)$

  +

  $P(Loc_0=\sim c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=\sim c)\ P(D_0=d|Loc_0=\sim c)\ P(Loc_1=\sim c|Loc_0=\sim c, Act_0=a)\ P(L_1=\sim l|Loc_1=\sim c)\ P(D_1=\sim d|Loc_1=\sim c)$
- [ ] $P(Loc_0=c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=c)\ P(D_0=d|Loc_0=c)\ P(Loc_1=\sim c|Loc_0=c, Act_0=a)\ P(L_1=\sim l|Loc_1=\sim c)\ P(D_1=\sim d|Loc_1=\sim c)$
- [ ] $P(Loc_0=c)\ P(Act_0=a)\ P(L_0=\sim l|Loc_0=c)\ P(D_0=d|Loc_0=c)\ P(Loc_1=c|Loc_0=c, Act_0=a)\ P(L_1=\sim l|Loc_1=c)\ P(D_1=\sim d|Loc_1=c)$

Question 9 **1 / 1 point**

Consider the following dynamic belief network from Poole and Mackworth, and consider the node in the lower right corner labeled "Cost paper$_3$" in particular:

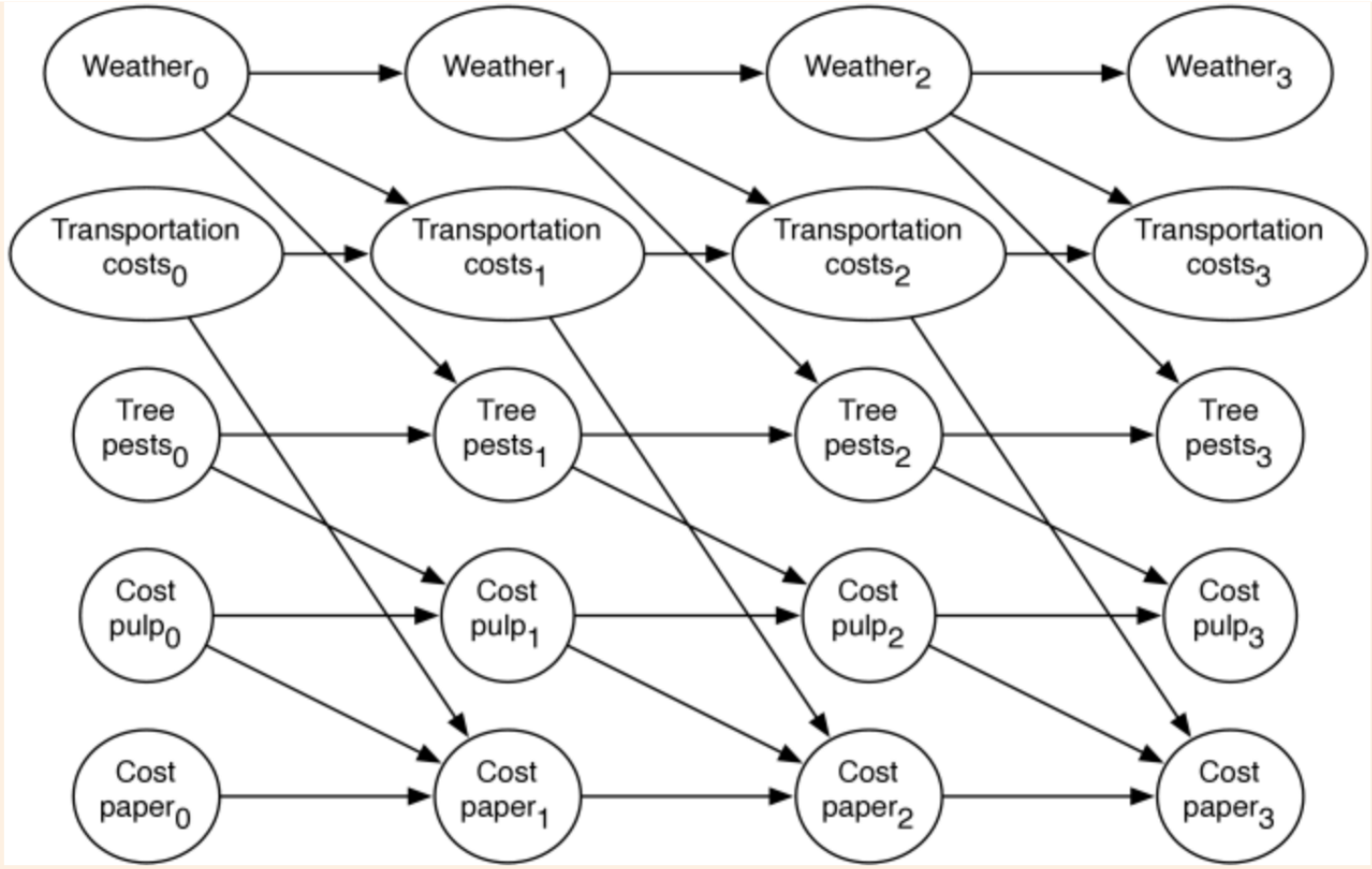


Figure 8.18: *Expanded dynamic belief network for paper pricing*

Which of the following options specify values that exist in the probability table for "Cost $paper_3$". Check all that apply.

- [ ] P(Cost $paper_3$|Transportation $costs_2$, Cost $pulp_2$, Cost $paper_2$)
- [ ] P(Cost $paper_2$|Transportation $costs_1$, Cost $pulp_1$, Cost $paper_1$)
- [ ] P(Cost $paper_1$|Transportation $costs_0$, Cost $pulp_0$, Cost $paper_0$)
- [ ] P(Cost $paper_3$|Transportation $costs_2$), P(Cost $paper_3$|Cost $pulp_2$), P(Cost $paper_3$|Cost $paper_2$)
- [ ] P(Cost $paper_3$|Transportation $costs_2$, Cost $pulp_2$, Cost $paper_2$, Weather$_2$, Tree $pests_2$)

Attempt Score: 100 %

Overall Grade (last attempt): 86.36 %

Done