HW6 Part A

Q1)

A) 16.5) a)

Solution:

- Network (i) implies that the wrapper color and shape are marginally independent which is not true.
- Network (ii) can represent any joint distribution as it is fully connected
- Network (iii) follows the generative story given in the problem
    1. The flavor is determined by which machine the candy is made by
    2. The shape is randomly cut
    3. The wrapper randomly chosen (independently of the previous).
- **Hence only Network (ii) and (iii) can represent the problem.**

B) 16.16 a) b)

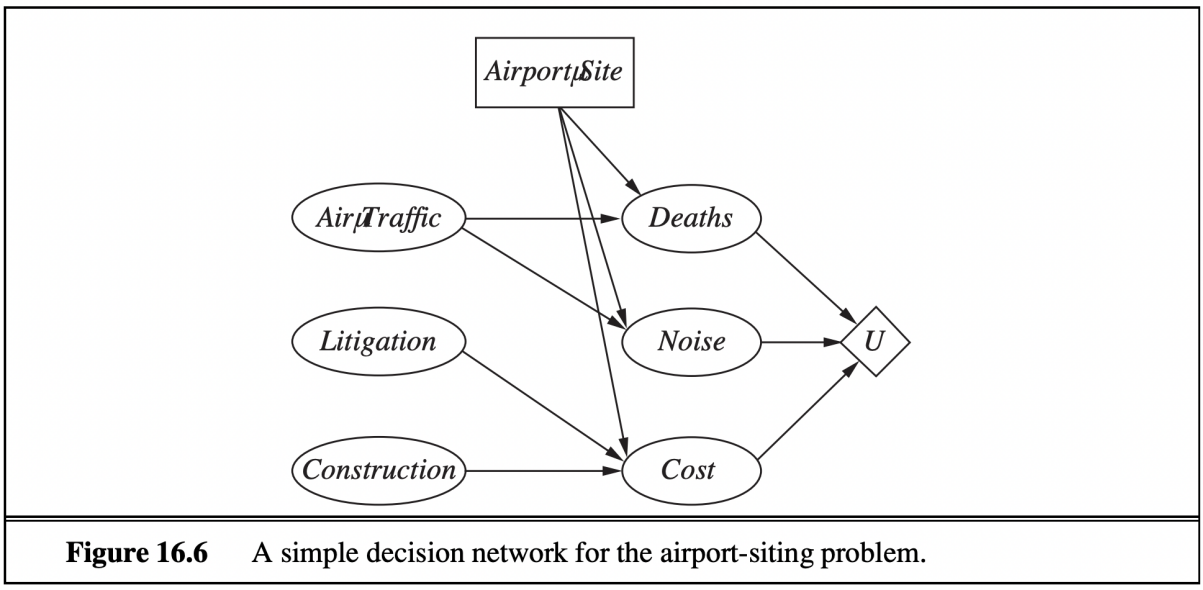


Figure 16.6 A simple decision network for the airport-siting problem.

Solution:

- The decision node `AirportSite` can take any one of the 3 sites as its decision A1, A2, A3
- The chance nodes are affected by their conditions.
- The `cost`, `deaths` and `noise` probabilities are influenced by `litigation`, `air traffic` and `construction` which can be considered as the hidden nodes and the Utility `U` depends upon `Cost`, `Death` and `Noise` values.
- The below diagram represents the state along with the utility value, `U`

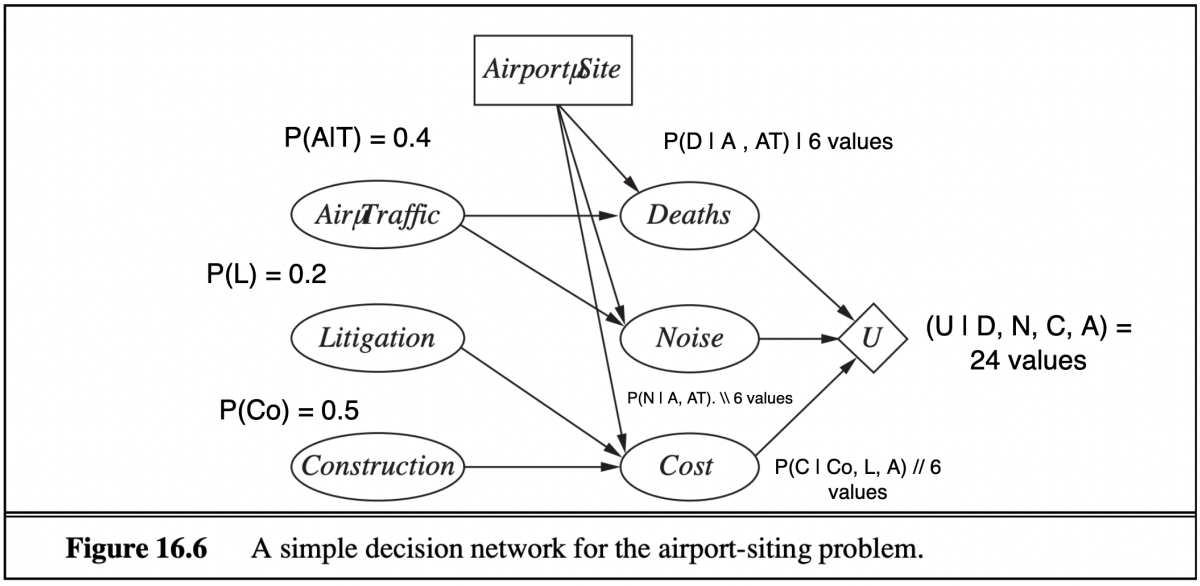


Figure 16.6 A simple decision network for the airport-siting problem.

- P(AT) → Probability of air traffic happening P(L) → probability of litigation
- P(C_0) → Probability of construction happening `AirportSite` → Decision node which can take 3 values `A1, A2, A3`
- P(D | A, AT) probability of death given Air Traffic and particular `AirportSite` (6 values)
- P(N | A, AT) probability of Noise given Air Traffic and particular `AirportSite` (6 Values)
- P(C | A,L,C0) probability if cost given a particular air traffic, construction and litigation (12 values)
- U | D N C A : Utility value given the `Death,` `Noise`, `cost` and decision of `AirportSite` (24 values)

With this data the Expected Utility and Maximum Expected Utility can be found using the
following formula, *(LHS : Maximum EU : RHS : Maximum of all EUS)*

Assume we have evidence `E=e` . Value if we act now:

`MEU(e) = max sum(P(s|e) U (s, a))`

Assume we see that `E' = e'`. Value if we act then:

`MEU(e, e') = max sum( P(s|e, e') U(s, a))`

BUT `E'` is a random variable whose value is unknown, so we don't know what `e'` will be

Expected value if `E'` is revealed and then we act:

`MEU(E, E') = sum(P(e|e) MEU(e, e'))`

---

2)

18.1

Solution:

The problem faced by an infant learning to speak and understand a language is a complex multivariate problem as it needs to recognize and produce speech, learn vocabulary and grammar, including the semantic and pragmatic interpretation of a speech act, etc. An infant learns through reinforcement learning aided by the environment it is set in. A point to note that the environment is just human - human interaction, but also a physical one where the child also uses its other senses like touch and sight to learn to speak.If the environment does not give necessary input then infant might fare poorly but if the environment is rich enough to provoke the infant to speak then the performance measure will be high. The child does learn through supervised learning when for example, adult says “shoe” or “belly button” while indicating the appropriate object. This is coupled with positive and negative feedback loops by the adults around the child help in classifications of the child’s attempted sentences.

18.3

Solution:

This will return a tree that is logically equivalent, assuming that the method for generating examples eventually generates all possible combinations of input attributes.

This is true because any two decision tree defined on the same set of attributes that agree on all possible examples are, by definition, logically equivalent.

18.17

Solution:

The examples map from [x1, x2] to [x1, x1x2] coordinates as follows:

Before Mapping

[x1 x2]	[x1x2]
1 1	1
1 -1	-1
-1 1	-1
-1 -1	1

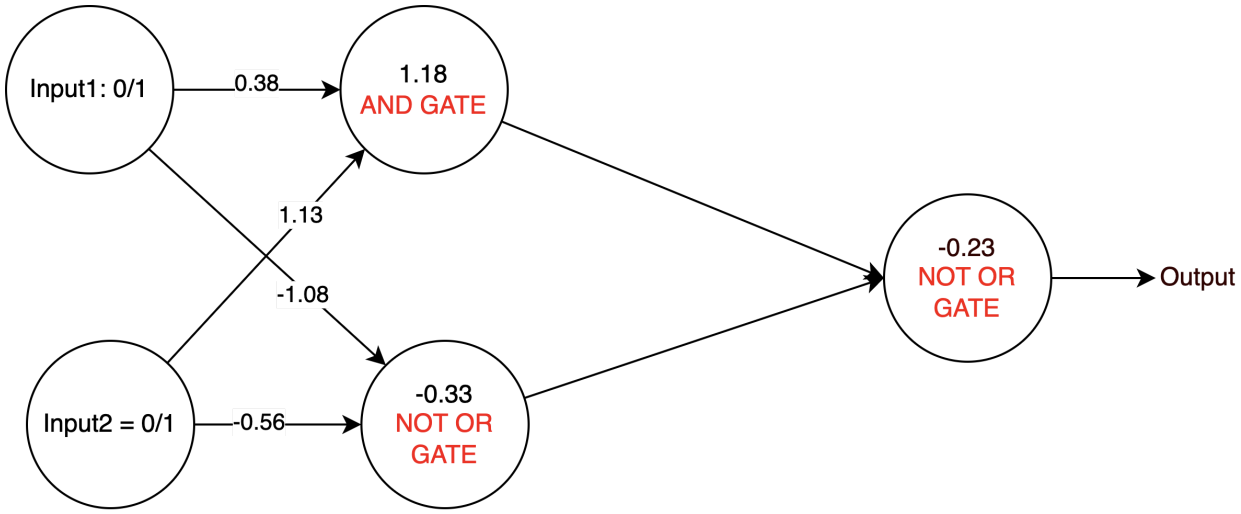
After Mapping

[x1, x2]	[x1 x1x2]
1 1	1 1
1 -1	1 -1
-1 1	-1 -1
-1 -1	-1 1

- The positive examples have x1, x2 = − 1 and the negative examples have x1, x2 = +1.
- The maximum margin separator is the line x1, x2 = 0, with a margin of 1.
- The separator corresponds to the x1 = 0 and x2 = 0 axes in the original space.

18.19

Solution:



Hidden Layer with 2 Neurons

3.

Solution:

- How to group users by their credit history?
  - SVM, K-means clustering
- What is the probability of not paying a loan?
  - The logistic regression
  - Random Forest
- How can we forecast credit score based on the usage of credit and previous loans?
  - RNN, Regression
- How can we determine which features influence paying a loan?
  - lasso regression, Ridge regression
- How can we increase the number of features using financial domain knowledge?
  - SVM