# CPSC 322, Practice Exercise
# Solutions to Sequential Decisions

## 1 Directed Questions

- How is a sequential decision problem different from a one-off decision problem? **Answer:** In a one-off decision problem, even if there are multiple decisions to make they can be treated as a single macro decision. That macro decision is made before any action is carried out. With a sequential decision problem, the agent makes observations, decides on an action, carries out the actions, makes some more observations in the resulting world, then makes more decisions conditioned on the new observations, etc.
- What types of variables are contained in a decision network? **Answer:** Chance nodes (random variables), decision nodes, and a utility node.
- What can arcs represent in a decision network? Relate this to the types of variables in the previous question. **Answer:** Arcs coming into decision nodes represent the information that will be available when the decision is made. Arcs coming into chance nodes represent probabilistic dependence. Arcs coming into the utility node represents what the utility depends on.
- What is a *no-forgetting decision network*? **Answer:** It is a decision network where the decision nodes are totally ordered and, if decision node $D_i$ is before $D_j$ in the total ordering, then $D_i$ is a parent of $D_j$, and any parent of $D_i$ is also a parent of $D_j$. This means that all the information available for the earlier decision is available for the later decision, and the earlier decision is part of the information available for the later decision.
- Define *decision function* and *policy*. **Answer:** A decision function for a decision variable is a function that specifies a value for the decision variable for each assignment of values to its parents. A policy consists of a decision function for each decision variable.
- A *possible world* specifies a value for every random variable and decision variable. Given a policy and a possible world, how do we know if the possible world satisfies the policy? **Answer:** The possible world satisfies the policy if the value for each decision variable in that possible world is the value selected in the decision function for that decision variable in the policy.
- To find an optimal policy, do we need to enumerate all of the policies? Why or why not? **Answer:** No, we can use variable elimination instead.

## 2 Sequential Decisions and Variable Elimination

Miranda is an enthusiastic gamer, spending quite a bit of time playing Wii video games and a fair amount of money buying them. She notices that her neighbourhood video store rents Wii games for much less than the cost of buying one. She realizes that renting the games might be a good way to test them out before she decides whether or not to buy them. Figure **??** represents her decision problem.

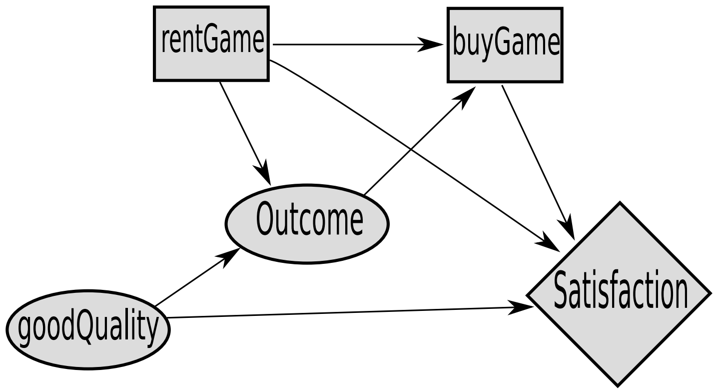


Figure 1: A decision problem.

Based on prior experience, Miranda expects that about 80% of video games will be good quality and the other 20% she won't care for. Based on her previous experiences renting video games, she also knows the following information:

$$P(Outcome = likesGame|goodQuality = True) = 0.85$$

$$P(Outcome = likesGame|goodQuality = False) = 0.10$$

The rental period is so short that it's not always possible to get a reliable estimate of whether the game is of good quality.

Below are the utilities for various outcomes of the decision process. You can think of the utilities as representing a combination of gaming enjoyment and money saved (*Satisfaction*).

| rentGame | buyGame | goodQuality | Satisfaction |
|---|---|---|---|
| T | T | True | 80.0 |
| T | T | False | -100.0 |
| T | F | True | 30.0 |
| T | F | False | -30.0 |
| F | T | True | 100.0 |
| F | T | False | -80.0 |
| F | F | True | 0.0 |
| F | F | False | 0.0 |

- If we carry out the variable elimination algorithm, what are the initial factors? **Answer:** There are 3 factors to begin with. f0(rentGame, buyGame, gameQuality) relates to the utilities. f1(rentGame, outcome, gameQuality) represents the probability of outcome given gameQuality and rentGame. f2(gameQuality) represents the prior for gameQuality.

- Which decision variable is eliminated first, and why?

  **Answer:** The decision variable *buyGame* is eliminated first because it is the last decision in the ordering.

- How is that decision eliminated? **Answer:** It is eliminated by choosing that values that maximize the utility. For example, if rentGame=T and outcome=Like, then buying the game results in a utility of 52.4 whereas not buying the game results in a utility of only 19.8. So we add a new decision function to our set of decision functions, specifying that when the parents have those assigned values, the decision is to buy the game. This is done for each combination of parent values.

- After that decision is eliminated, which variable is eliminated next, and why? **Answer:** The random variable *outcome* is eliminated next, because it is no longer parent to any decision variable (since we removed *buyGame*).
- What is the optimal policy for this decision problem? **Answer:** The optimal decision for rental is not to rent.

  The optimal decision for buying is to buy the game in every case except where the game was rented and she disliked it.
- What is the expected utility of following the optimal policy? **Answer:** The expected utility of following that optimal policy is 64.0.

Use the AISpace decision applet to represent and solve this decision problem, and to check your answers. The representation we have used is in file `wii.xml`.

## 3 Learning Goals

You can:

- Represent sequential decision problems as decision networks and explain the no-forgetting property.
- Verify whether a possible world satisfies a policy and define the expected value of a policy.
- Compute the number of policies for a decision problem.
- Compute the optimal policy by Variable Elimination.