

Decision Theory Study Guide

Artificial Intelligence

1 Preferences

1. Consider the following preferences: an agent prefer puppies to kittens, kittens to babies, and babies to puppies. Are these preferences rational? Why or why not?

2 Utilities

1. I prefer iced sugar cookies to chocolate chip cookies. If I were to assign numerical utility values to each of these cookies I would assign a utility of 12 to chocolate chip cookies and 10 to iced sugar cookies. Are my utility values valid? Why or why not?

2. Given two attributes of cookies, *icing* and *nuts*, an agent has the following preferences:

$$icing \succ \neg icing$$

$$\neg nuts \succ nuts$$

What can you say about the agent's choice between cookie *A* with icing and no nuts, and cookie *B* with nuts and no icing?