Advanced AI Techniques

Prof. Dr. Burgard, Prof. Dr. Nebel, Dr. Kersting
M. Ragni, A. Rottmann
WS 2006/2007

University of Freiburg
Department of Computer Science

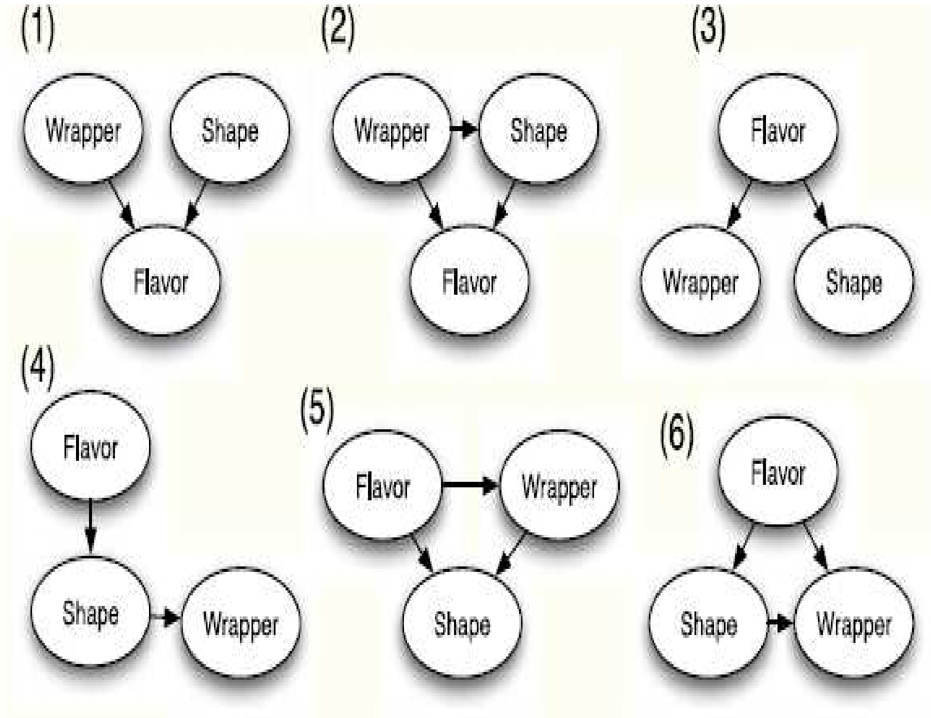
Exercise Sheet 3

Due: Tuesday, 21. November 2006

Exercise 3.1 (Bayesian Networks)

The monitoring of the exercises indicates that there are still problems in dealing with Bayesian Networks. Therefore we provide a last exercise.

The Surprise Candy Company makes candy in two flavors: 30% are Anchovy (which smell of the salty sea) and 70% are Berry (which smell delightful). All candies start out round and look the same. Someone (who can smell) trims some of the candies so that they are square. Then, a second person who can not smell wraps each candy in a red or brown wrapper. 80% of Berry candies are round and 74% have red wrappers. 90% of Anchovy candies are square and 82% have brown wrappers. All candies are sold individually in sealed, identical, black boxes! You have just bought a box.



(a) Which network(s) can correctly represent $P(Flavor, Wrapper, Shape)$?

(b) Which network(s) assert(s)$P(Wrapper|Shape) = P(Wrapper)$?

(c) Which network(s) assert(s) $P(Wrapper|Shape, Flavor) = P(Wrapper|Shape)$?

(d) From the problem description, what independence relationships should hold for this problem? Which network is the best representation of this problem?

(e) What is the probability that the candy has a red wrapper?

(f) In the box is a round candy with a red wrapper. What is the probability that it is a Berry candy?

(g) If you tried to model the problem with network (2), can you give $P(Wrapper = Red|Shape = Round)$? If so, give the answer.

Exercise 3.2 (Binary bandit problem)

Assume you want to give a nice present to your uncle. You have the choice to either give him an aftershave (action $a = A$) or a book (action $b = B$). If he is happy with the present, he will smile (this make you happy: reward $r = success$), otherwise frown (reward $r = failure$).
Implement the supervised algorithm for learning (over the next 50 years) whether it is better to give him an aftershave or a book as present. Assume that the uncle's taste does not change, i.e. the success probabilities $p_A = P(r_t = success \mid A)$ and $p_B = P(r_t = success \mid B)$ are independent of the time point t.
Trigger 100 runs (each run simulating 50 plays) for each of the following binary bandit tasks and determine (for each task) in how many cases the intuitively better action was "learned" as preferred action.

(a) $p_A = 0.1, p_B = 0.8$

(b) $p_A = 0.2, p_B = 0.1$

(c) $p_A = 0.8, p_B = 0.9$

Compare this results with a reinforcement learning approach. For this implement an ϵ-greedy action selection algorithm with $\epsilon = 0.3$ and repeat the same experiment using this function. What happens if you decrease ϵ to 0.01?

Please hand in a joint solution of three students and write all names on the sheet.