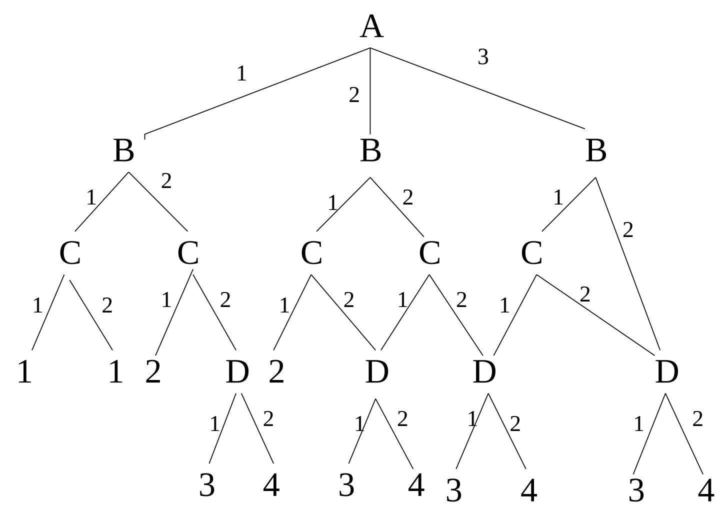
CS472 / CM472 / CS672 - Fall 2006: Assignment 4

Instructor: Pascal Poupart   Office: DC2514   Email: ppoupart@cs.uwaterloo.ca
Classroom: RCH305   MWF 4:00-5:20   Office Hours: Mon, 10-11
Web Site: www.student.cs.uwaterloo.ca/ ~cs472/

Due: Tuesday, December 5, (at the beginning of class)

1. ADD canonical form

   (a) (5 marks) Reduce the following ADD to its canonical form (assuming ordering A, B, C, D). At each step indicate which reduction rule is applied.



   (b) (5 marks) Prove by induction that the reduction algorithm always yields the smallest ADD in terms of the number of nodes and edges (for a given variable ordering).

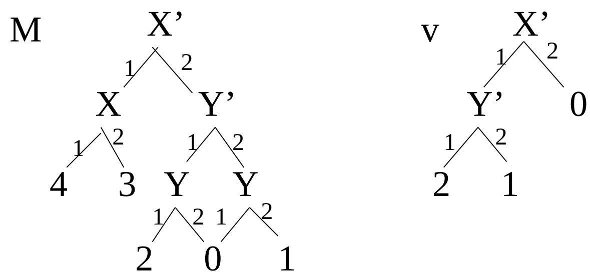
   (c) (5 marks) Assuming that verifying the equivalence of two reduced ADDs can be done in constant time, show that the reduction algorithm takes time linear with respect to the number of nodes and edges of the initial ADD.

   (d) **Bonus question** (3 marks) Explain how to verify the equivalence of two reduced ADDs in constant time.

   (e) **Graduate student question** (5 marks) Prove that the canonical ADD obtained by the reduction algorithm is unique.

2. (10 marks) Matrix operations with ADDs

   Let X, Y, X', Y' be binary variables. Let M be a 4x4 matrix with rows indexed by X and Y, and columns indexed by X' and Y'. Let v be a column vector indexed by X', Y'. The ADD representation of M and v are:



   Compute Mv. Show your steps (i.e., show the intermediate ADDs generated as additions and multiplications are performed).

3. Prove that policy iteration converges to an optimal policy.

   (a) (15 marks) Prove by induction that the policy improvement step always generates a policy at least as good as the current policy.

      i. (5 marks) Base case.
      Let $V^{\pi}(s)$ be the value of policy π at state s. Let $Q^{\pi_1}_{\pi_2}(s)$ be the value of executing π_1 for one step followed by π_2 for all following steps.

      $$Q^{\pi_1}_{\pi_2}(s) = R(s) + \gamma \sum_{s'} \Pr(s'|s, \pi_1(s))V^{\pi_2}(s')$$

      Let $\pi'(s) = argmax_a R(s) + \gamma \sum_{s'} \Pr(s'|s,a)V^{\pi}(s')$, then show that $Q^{\pi'}_{\pi}(s) \geq Q^{\pi}_{\pi}(s)\ \forall s$.

      ii. (5 marks) Induction step.
      Suppose that $Q^{n,\pi_1}_{\pi_2}(s)$ is the value of executing π_1 for n steps followed by π_2 for the following steps.

      $$Q^{n,\pi_1}_{\pi_2}(s) = R(s) + \gamma \sum_{s'} \Pr(s'|s, \pi_1(s))Q^{n-1,\pi_1}_{\pi_2}(s')$$

      Suppose that $Q^{n,\pi'}_{\pi}(s) \geq Q^{n,\pi}_{\pi}(s)$ and that π' is defined as in 3(a)i, show that $Q^{n+1,\pi'}_{\pi}(s) \geq Q^{n+1,\pi}_{\pi}(s)\ \forall s$.

      iii. (5 marks) Show that as $n \to \infty$, $Q^{n,\pi'}_{\pi}(s) \to V^{\pi'}(s)$ and conclude that policy improvement always generates a policy at least as high as the current policy.

   (b) **Graduate student question** (3 marks) Show that policy iteration converges in a finite number of steps.

   (c) **Bonus question** (3 marks) Show that policy iteration won't get stuck in a local optimum (i.e., it is guaranteed to converge to the optimal policy).