There are no tutorials on December 5 and 6.

Exercise 1: (4 Points)
We generalize the cardinality-robust version of Vertex Cover as follows. Partition the set of possible edges in the graph into sets $B_1, \ldots, B_\ell$. The scenario set $\mathcal{E}$ contains all sets $E$ for which $|E \cap B_i| = k_i$. Extend the approximation algorithm from the lecture and outline its analysis.

Exercise 2: (1+1+1 Points)
We consider a Markov decision process with $S = \{1, 2, 3\}$, $A = \{a, b\}$. The state transitions are deterministic as displayed in this diagram; the numbers in the edge labels are the respective rewards.

We consider an infinite time horizon with discount factor $\gamma = \frac{1}{2}$.

(a) Give an optimal policy and the function $s \mapsto V^*(s)$.

(b) Perform the first six steps of value iteration starting from $W(0) = (0, 0, 0)$.

(c) Perform policy iteration until convergence starting from the policy that always uses action $a$.

Exercise 3: (4 Points)
We define a more cautious version of value iteration. It uses the operator $T'$, which is defined by $T'(W) = \eta T(W) + (1 - \eta)W$ for an arbitrary $\eta \in (0, 1)$. Show that this algorithm also converges to the unique fixed point of $T$.

Exercise 4 on the next page.
Exercise 4: (3+2+2+2 Points)
For the following single-armed bandits, give the fair charges of all states. Unless states otherwise, the transitions are deterministic. Justify your statements if necessary. For part (a), consider $\gamma = \frac{1}{2}$; for the remaining parts an arbitrary $\gamma \in (0, 1)$.

(a) rew: 3 rew: 12 rew: 3 rew: 3 rew: 0

(b) rew: 0 rew: 1

(c) rew: 1, prob: 1/2 rew: 0, prob: 1
rew: 1, prob: 1/2

(d) rew: 0, prob: 1/2 rew: 1, prob: 1
rew: 0, prob: 1/2