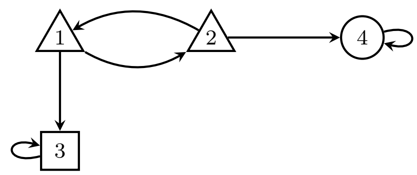
1. In the following Markov Chain, find the probability to reach 1-sink from each node:



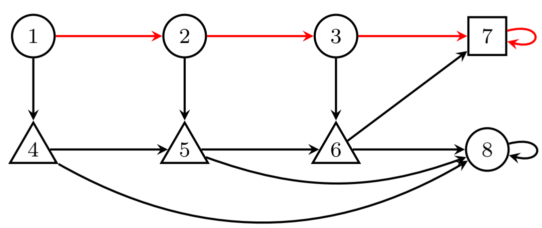

   Node 3 is the 0-sink and node 4 is the 1-sink. Triangle shaped nodes are average vertices - each outgoing edge from a triangle node has $\frac{1}{2}$ probability.

2. Consider a Markov Chain $G$ with states $\{1, 2, \ldots, n-1, n\}$, with $n-1$ and $n$ being sink states. Assume that $G$ is stopping, that is, from every state there is a non-zero probability to reach either $n-1$ or $n$. Let $\bar{\lambda}$ be the vector denoting probabilities to reach $n$ from each vertex.

   Consider a modified Markov Chain $G'$: its states are $\{1, 2, \ldots, n, 1', 2', \ldots, (n-2)'\}$; the edges from unprimed states $\{1, \ldots, n\}$ are the same as in $G$; edges for primed states are as follows - if vertex $i$ has edges to $j, k$ with $j \geq k$, then $i'$ has edges $i' \to j$ and $i' \to k'$ (essentially, one edge goes to unprimed copy, and one edge goes to primed copy, and if $j = k$, both go to unprimed copy). If $k$ equals $n-1$ or $n-2$, then $i' \to k$, instead of $i' \to k'$.

   Find the probabilities to reach $n$ in $G'$, in terms of $\bar{\lambda}$.

3. Run the strategy improvement algorithm for the following MDP, starting from the strategy marked in red. Node 8 and 7 are the 1-sink and 0-sink respectively.




4. Run value iteration for the above MDP. What is the LP corresponding to the above MDP?

5. Consider the following modification of the strategy improvement algorithm:

**algorithm** $modified-strategy-improvement(G)$
  $\sigma \leftarrow$ an arbitrary positional strategy
  $v_\sigma \leftarrow$ probabilities to reach 1−sink in $G_\sigma$
  **repeat**
    **for every** node $i \in V_{max}$
      $\sigma'(i) := \arg\max\{v_\sigma(j), v_\sigma(k)\}$ where $j$ and $k$ are children of $i$
    $\sigma \leftarrow \sigma'$
    $v_\sigma \leftarrow$ probabilities to reach 1−sink in $G_\sigma$
  **until** $\sigma$ is optimal

   In the normal strategy improvement algorithm, during each iteration only one node $i \in V_{max}$ is picked and the strategy is made to point to the child with bigger value. In the above algorithm, every node is made to point to the maximum child.

   Assume that $G$ is a stopping MDP. Is the above algorithm correct? Justify.