1 Source-Set Selection Problem

The Markov Decision Problem formulation is as follows $M \triangleq (S, A, T, R, \gamma)$:

- States S: A binary vector where the i-th element is 1 iff x_i is included in the source set.
- Actions A: Including or removing a node from the source set.
- Reward function: Quantifies the skew of the PD-tree based on the source set $R(s, a, s')$
- Transition function $T : S \times A \times S \to [0, 1]$

Observe, act and receive feedback. The agent's goal is to learn a policy that lets the agent accumulates the highest possible reward over time.