Agent

1. Num of Translation String
2. Initial Num of Candidates = 3
3. Min Majority = 5 => Num of Votes = Min\_Maj + (Min\_Maj - 1) \* (Num\_Candidate - 1)
4. SLA = 24hrs
5. Budget = $100
6. Tolerance Rate = [0, 1]
7. Cost = [$0.01, $0.1]

Simulator

1. Time
   1. Work Time => Pareto
   2. Response Time => Gaussian
   3. Total Time => ~Gaussian
2. Cost (X 1 / Response Time
3. Work Time should fall in a certain range: Work Time (X [min, max]
4. Failure Rate = [0, 1] - 1- Success Rate
   1. NOTA
   2. Non-Majority
      1. Difference Threshold: X1 – X2 = Ceiling(Num\_Votes / Num\_Candidates)

Positive Rewards:

R = {[1/ max(Response\_Time, min\_RT) - (Work\_Time - lowerBound)(Work\_Time - UpperBound)]\*(Success\_Rate)}

Negative Rewards:

R = - {[max(Response\_Time, min\_RT) + (Work\_Time - lowerBound)(Work\_Time - UpperBound)]\*(-Success\_Rate)}

Action Space:

Constituency: Same/Different = [0.8, 1]

Candidates: Eliminate least voted / Same candidate / Additional Candidate = [1, 0.8, 1.2]

Voting: Single-out / Instant-runoff (Preferential) = [0.8, 1]

Cost: [$0.01, $0.1]

Action Dimension:

[(history\_size, 4), ...]

4 dimensions: Constituency, Candidates, Voting, Cost

Translation:

Cost: [$0.1]

Time: 1 hour

Observation dimension:

1. Response Time
2. Work Time
3. Success Rate

Initial Round: Cost -> Voting -> isTerminal?

Following Rounds: Constituency -> Candidates -> (if additional candidate -> cost -> collect translation) -> Cost -> Voting -> isTerminal?

If Same constituency + Same Candidate, Success Rate = 0

Relationship between Actions and Rewards:

Response Time : Gaussian(mu = 1/cost, Fixed std) \* 1.2 (if different constituency)

Work Time: lower bound for one candidate = 3 seconds. (Pareto (Fixed shape) + 1 )\* (3 \* Num\_candidates)

Upper bound = 10 \* Lower bound = 10 \* (3 \* Num\_candidates)

Translation time = 1hr (Fixed for now)

Success Rate:

Constituency: Same/Different = [0.8, 1]

Candidates: Eliminate least voted / Same candidate / Additional Candidate = [1, 0.8, 1.2]

Voting: Single-out / Instant-runoff (Preferential) = [0.8, 1]

Cost: [$0.01, $0.1]

If Same constituency + Same Candidate, Success Rate = 0

Agent => action => Simulator => Sample on Distribution => (ResponseTime, WorkTime, Success Rate) => Reward => Agent => Update Q-table/DQN