

i MDP

You wish to develop a (phone-based) spoken dialogue system that will call random U.S. citizens in order to collect opinion poll data for the next US election, with two candidates on the ballot: Kamala Harris and Ivanka Trump.

This system is framed as a *Markov Decision Process* (MDP) formalised as such:

- We have five possible states:
 - s_1 is the starting state
 - s_2 if the callee indicated their intention to vote for Kamala Harris
 - s_3 if the callee indicated their intention to vote for Ivanka Trump
 - s_4 if the callee expressed something else (that was not understood)
- The set of actions that can be taken by the dialogue system are as follows:
 - a_1 : Say "Hi, I'm a automated bot developed to collect polling data. May I ask you for whom you plan to vote in the next election?"
 - a_2 : Say "Sorry I did not understand. Who do you wish to vote for?"
 - a_3 : Say "Ok, thank you for your help, and have a nice day!"
- The transition model is as follows:
 - In state s_1 , only action a_1 is possible, with three possible transitions:

$$P(s' = s_2 | s = s_1, a = a_1) = 0.48 P(s' = s_3 | s = s_1, a = a_1) = 0.40 P(s' = s_4 | s = s_1, a = a_1) = 0.12$$
 - In states s_2 and s_3 , only a_3 is possible and terminates the dialogue.
 - In state s_4 , only a_2 is possible, with the following transitions:

$$P(s' = s_2 | s = s_4, a = a_2) = 0.36 P(s' = s_3 | s = s_4, a = a_2) = 0.32 P(s' = s_4 | s = s_4, a = a_2) = 0.32$$
- Finally, the reward model is defined as such:
 - $R(s = s_2, a = a_3) = R(s = s_3, a = a_3) = 10$ (if the system manages to register the callee's political preference)
 - $R(s = s_4, a = a_2) = -1$ (to capture the annoyance of asking the callee to repeat)
 - Other actions have a reward of zero.

4(c) MDP (a)

Based on this MDP model, calculate the expected cumulative reward of asking the callee to repeat when their answer was not properly understood, that is: $Q(s=s_4, a=a_2)$. You can assume a discount factor of 0.9.

Fill in your answer here

Maximum marks: 8

4(d) MDP (b)

One limitation of this MDP model is that it assumes that the dialogue system will always be 100 % certain it has correctly understood the political preference expressed by the callee. In practice, this will not always be the case, because of e.g. speech recognition or NLU errors, or because the callee may intentionally provide unclear or misleading information. How could this model be modified to capture those uncertainties?

Fill in your answer here

Maximum marks: 3