CS534 ARTIFICIAL INTELLIGENCE

Project 4 Document

Yunhe Tang

QUESTIONS:

1. **Chapter 15 #15.2**

   a. If it is converged at a fixed point, then we must have $P(R_t|u_{1:t}) = P(R_{t-1}|u_{1:t-1})$. According to the formula, let p is this fixed point:

   $$p = \frac{(p\times0.7+(1-p)\times0.3)\times0.9}{(p\times0.7+(1-p)\times0.3)\times0.9 + (p\times0.3+(1-p)\times0.7)\times0.2}$$

   The result we get from solving the function is p = 0.893.

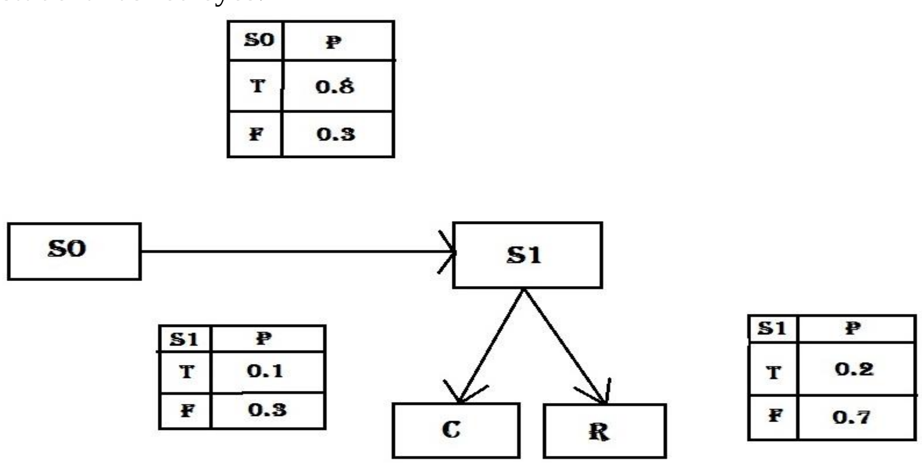
   b. If probabilities converges only giving two observations, then the probability replies only on transition model. Therefore, we have:

   $$P(r_{2+k}|U_1,U_2) = 0.7 \times P(r_{2+k-1}|U_1,U_2) + 0.3 \times (1- P(r_{2+k-1}|U_1,U_2))$$

   Since convergence, $P(r_{2+k}|U_1,U_2)$ should equips to $P(r_{2+k-1}|U_1,U_2)$, therefore, from solving this equation, we get $P(r_{2+k}|U_1,U_2) = 0.5$

2. **Chapter 15 #15.13**

   a. The Dynamic Bayesian Network is as below: C represents sleep in class, R represents student has red eyes.



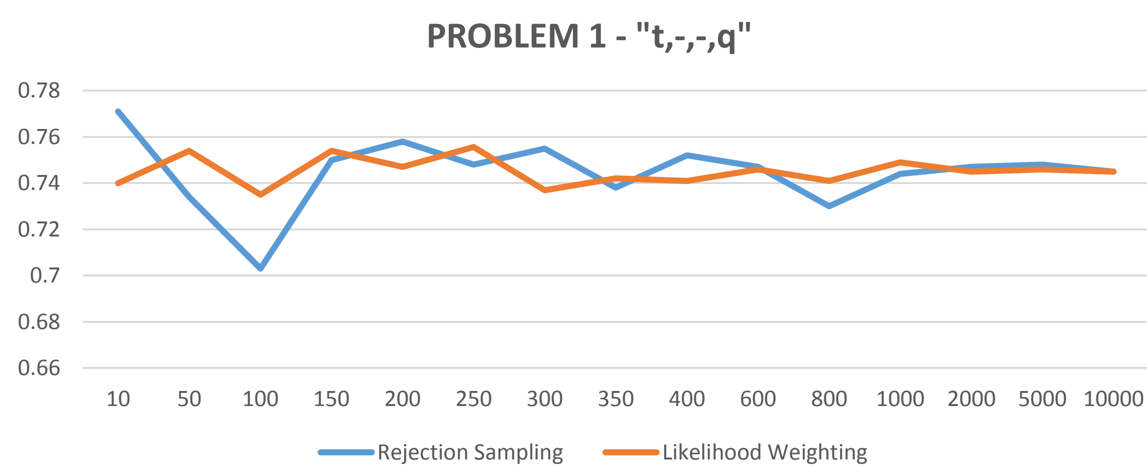
b. The probability table in Hidden Markov Model is as below:

R	C	P
T	T	0.02
T	F	0.18
F	T	0.08
F	F	0.72

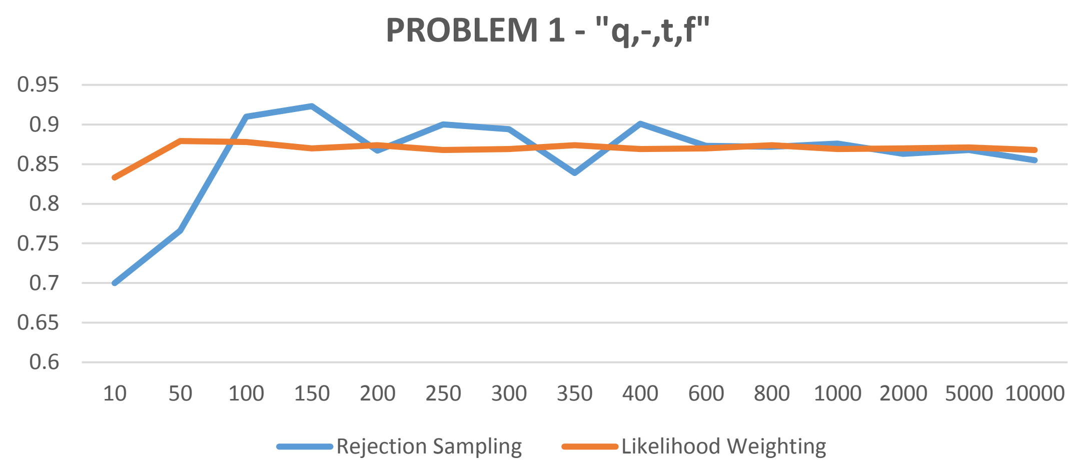
IMPLEMENTATION:

Project 1: Inference

1. **Average the results of the 10 runs for each algorithm for each case.**
   For the first input: *t,-,-,q*, the estimate results are graphed in following:



For the second input: *q,-,t,f*, the estimate results are graphed in following:



2. **Was there any difference in the convergence rate? If so, for each case, state which algorithm converged faster and explain why.**

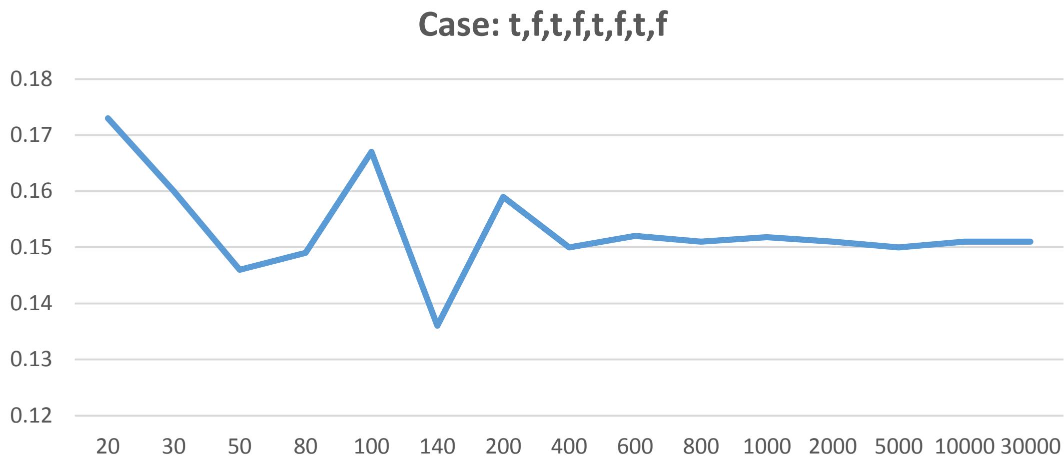
   It is observable that Likelihood Weighting converge faster than Rejection Sampling. For the first input, result of Likelihood Weighting converges when the number of samples reaches roughly 300, while Rejection Sampling converges at about 1000. For the second input, Likelihood Weighting converges with very few samples, while Rejection Sampling converges after 600.

   The reason why Likelihood Weighting converge faster than Rejection Sampling is that a lot of samples used in Rejection Sampling are not calculated. Because in the last phase of Rejection Sampling, samples that fail to match the evidence are throw away. In another word, part of samples are wasted. In contrast, every sample in Likelihood Weighting is fully used. Since each sample matches the given evidences. In conclusion, fewer samples can achieve the same level of accuracy as Rejection Sampling gives.

Project 2: Filtering

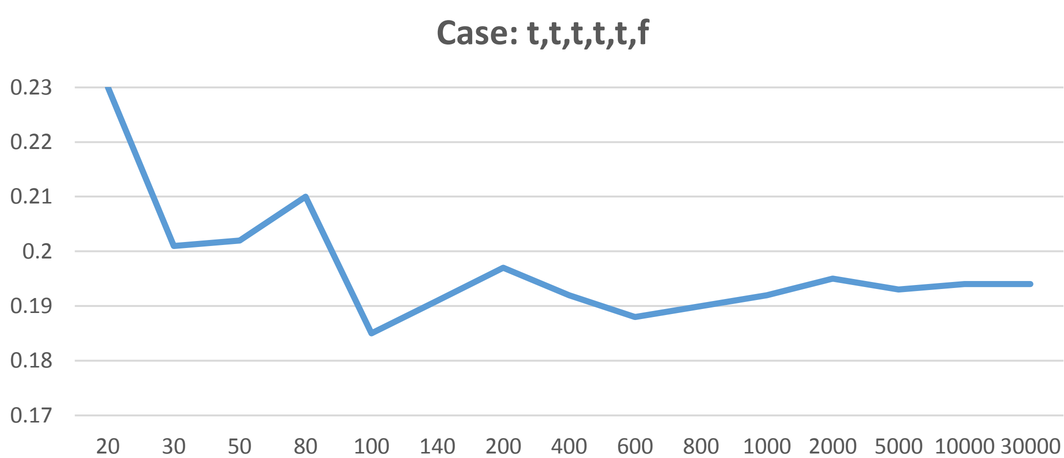
1. **Average the results of the 30 runs for each number of particles and plot them in one graph per case.**

For the first input: *t,f,t,f,t,f,t,f*, the result is shown below:



For the second input*: t,t,t,t,f,f,f,f*, the result is shown below:

For the third input: *t,t,t,t,t,f,* the result is shown below:



2. **For which of the three cases does the particle filter perform best?**

   From graphs presented above, we can observe that filter perform best on the second input. The reason is that the evidence is more consistent than the other two, which different evidence appear alternatively.

   We can see that there are continuously four ‘f’ in the tail of string *“ttttffff”*, it is therefore the possibility of raining is reduced repeatedly, which eliminates the uncertainty. The first one obviously contains large amount of uncertainty, since the alternative appearance of ‘t’ and ‘f’ do no good to stabilize the possibility distribution. The analysis for the third input is in below.

3. **Why is producing a good estimate for the third case particularly unlikely with a small number of particles?**

   From observation we can tell that the last observable evidence is ‘f’, which is different than any of evidence precede itself. There are two situations can potentially cause this result: day six is raining but he forgot to take the umbrella, and day six stop raining. Since there is no following evidence can be taken advantage, the algorithm has no way to give a good estimation at day six.

   One way to improve algorithm is to introduce randomness into the sampling process. Instead of relying too much on the observable evidence, we only select, say 80% of the samples from the original samples when doing the resampling. By doing so, 20% of the samples are kept without weighted selected according to the evidence. In this sense, algorithm trust less on the “confusing” evidences and rely more on transition model, increasing stability.