CS 228T QUIZ 6

1. Briefly explain how one might use Markov chain Monte Carlo to approximate network feature queries or the Bayesian estimation rule. Specify which distribution we need to sample from, what the state space of the chain is, and what kinds of moves we might consider in that space. What property of the moves suggested in the reading allows us to compute the acceptance probability particularly efficiently?

2. This question is about structure learning.

(a) What is the Bayesian score?
(b) What is the Cheeseman-Stutz score, and why is it useful?

3. Briefly explain the meaning of the term 'nonparametric'. Is k-nearest neighbors a (not necessarily Bayesian) nonparametric method?

4. What does it mean for a distribution over finitely many variables to be exchangeable? What does it mean for a stochastic process to be (infinitely) exchangeable?

5. What is the main benefit of using a Dirichlet process mixture model instead of a standard finite mixture model (potentially with a Dirichlet prior)?

6. Consider balls $x_1, \ldots, x_n$ drawn from a Pólya urn (with $\alpha_0 = 1$). What is the probability that x_i and x_j have the same color?

7. This question references the paper on split-merge MCMC for the conjugate Dirichlet process mixture model. Provide very brief answers to the following questions.

(a) What is the potential problem with using Gibbs sampling for the conjugate Dirichlet process mixture model that the split-merge method addresses?
(b) Explain why both the integrals in equation 4 can be computed analytically when G_0 is conjugate to F.
(c) How does the paper justify being able to choose the launch state at random?
(d) Explain the role of (restricted) Gibbs sampling as a 'subroutine' in the split-merge algorithm.

Date: September 29, 2020.