Students JHID:

Grader: Can Zhao

	Full points	Your points	Comments
2.1	50	26 + 24*0.8 = 45.2	Second submission
2.2	30	30	
2.3(a)	5	5	
2.3(b)	15	12	Need to explain how to compute $P(X_i, Pa_i|o[m], \theta^t)$
Total	100	91	Excellent!

Grading notes:

	Full points	If final result is not correct, then how to give points:
2.1	50	Depends on the fraction of correct tests. If only correct for the given simple network, only get 25pts. If failed for one of the test network, get 36 pts
2.2	30	Depends on the fraction of correct tests
2.3(a)	5	Total 13 parameters. If partly correct,1-4 pts.
2.3(b)	15	**Comments from Dr. Saria:** EM in the E-step computes the *expected sufficient statistics*. See my lecture slides. For example, in a Bayesian network with CPDs that are multinomials, the sufficient statistics are the counts. And, the E-step would compute the expected counts. Similarly, in a mixture of Gaussian model, the sufficient statistics are is $\sum x_i$ and $\sum x_i^2$ and the E-step computes the expected value of these sufficient statistics. To compute the expected value, for each individual sample, it has to do inference. For example, in an HMM, we compute the probability distribution over the hidden state variables given the observed data for each sample. **Grading:** In E-step, should estimate expected sufficient statistics.(See Module 6, Page 28) If students estimated missing data values, it's wrong, 8 pts off. If students estimated expected sufficient statistics, but didn't give how to compute $P(X_i, Pa_i|o[m], \theta^t)$, 3 pts off. See Koller's 19.2.2.2