CPSC 633 Final (4/27/2009, Mon)[1]

Last name: _____, First name: _____

Time: **9:10am-10:00am (50 minutes)**, Total Points: **100**

Subject	Score
Genetic Algorithms	/20
Hypothesis Testing	/20
Bayesian Learning	/20
Computational Learning Theory	/20
Instance-based Learning	/20
Total	/120
Scaled final score	min(,100)

- The total adds up to 120. You may solve all problems, or solve problems that add up to 100 points. In case your score is above 100, your final score will be 100. This will allow you to skip some problems that you are not confident about.
- Total 6 pages (10 questions), including the cover sheet.
- Be as **succinct** as possible. Usually, one sentence is enough to answer one specific question.
- Read the questions carefully to see what kind of answer is expected (*explain blah* **in terms of ...** *blah*).
- If you feel that the question is not specific enough, please do ask.
- This is a closed-book, open-note (the two-sheet note you brought) exam.

[1] Instructor: Yoonsuck Choe.

1 Genetic Algorithms

Question 1 (12 pts): What are the three factors determining the expected value $E[m(s, t+1)]$ (where $m(s, t+1)$ is the number of individuals representing schema s at time t) and explain how each factor contributes to $E[m(s, t+1)]$. **Hint:** Think about the typical ingredients of GA.

Question 2 (12 pts): In neuroevolution, there is the problem of "competing convention". Explain what this is and why it can cause a problem, especially when crossing over. Hint: think about permutation of the hidden layer units.

2 Hypothesis Testing

Question 3 (12 pts): Explain why the average error for all instances can be seen as a random variable, and how that fact relates to the confidence interval of the estimated error rate. Hint: Error made on a single training instance can itself be treated as a random variable.

Question 4 (12 pts): Why is it the case that the confidence interval reduces if you have more samples n?

3 Bayesian Learning

Question 5 (12 pts): (1) Show how h_{MAP} can be the same as h_{MDL}, and (2) explain why this is a significant result.

Question 6 (12 pts): Does h_{MAP} give the same result as an optimal classifier? Explain why or why not.

4 Computational Learning Theory

Question 7 (12 pts): A particular hypothesis space H having a high VC dimension has both advantage(s) and disadvantage(s). What are those advantage(s)/disadvantage(s)? Explain in terms of (1) richness/expressibility of the hypotheses and (2) sample complexity.

Question 8 (12 pts): Explain why the mistake bound of the Halving Algorithm is $\lfloor \log_2 |H| \rfloor$. **Hint:** The classification is incorrect if a majority of the hypotheses (i.e., more than half) in the current VS voted for the wrong answer, and initially, $|VS| = |H|$.

5 Instance Based Learning

Question 9 (12 pts): Explain the difference between lazy learners and eager learners.

Question 10 (12 pts): How does the Radial Basis Function network utilize the training data, in the style of instance based learning?