

This assignment covers material from Modules 3 and 4 (part of which will be covered tomorrow). It is due on **Thu, 19 June 2008, 18:00** (no late hand-ins, please!). Please submit your solution as a PDF file (which may be obtained by scanning handwritten pages) via e-mail to [hoos@cs.ubc.ca](mailto:hoos@cs.ubc.ca).

Feel free to discuss the problems and solution ideas with other students, but you need to work out and write down the actual solutions on your own.

Keep all answers as concise as possible — all else being equal, short and precise answers will be scored higher.

**Problem 1 (5+10=15 marks)**

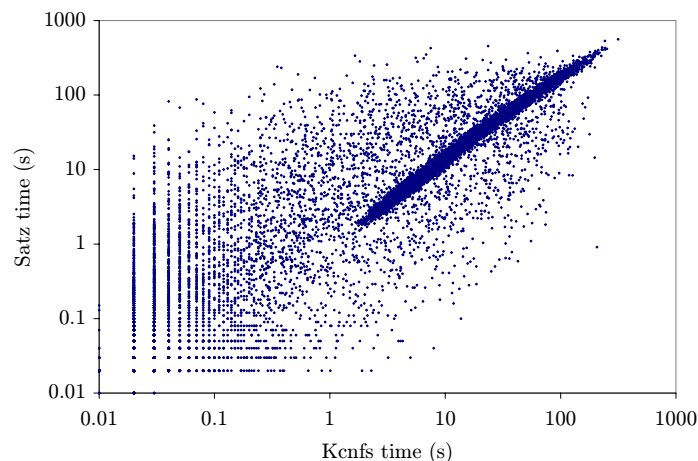
- (a) Briefly explain the differences between an SCD and an RTD.
- (b) Briefly outline the procedure you would use for empirically comparing the performance of two Las Vegas algorithms on a set of benchmark instances. (Your answer should be no longer than one paragraph.)

**Problem 2 (10 marks)**

Briefly describe the concept of parallelisation by means of multiple independent runs and explain how for a given Las Vegas algorithm  $A$  the effectiveness of this parallelisation approach is related to the shape of  $A$ 's RTDs. (Your answer should be no longer than one paragraph.)

**Problem 3 (6+6+8=20 marks)**

Consider the runtime measurements for Kcnfs and Satz, two SAT algorithms, on a large set of benchmark instances, illustrated in the following scatter plot:



- (a) What do you notice in this plot? There are at least three observations you can make, one of which was a surprise to the researchers who produced this data.
- (b) Briefly discuss possible explanations for your observations.
- (c) Briefly discuss how you could further investigate your proposed explanations.

**Problem 4 (5+10+5=20 marks)**

Download the RTD data set 'ils-lin318-opt-rtd.dat' and 'mmas-lin318-opt-rtd.dat' from the course home page and perform the following analysis using R (<http://www.r-project.org/>) and/or Gnuplot (<http://www.gnuplot.info/>). The second column in each of the two data sets represents the CPU times measured in multiple independent runs of two different algorithms (ILS and MMAS) applied to the same instance of the Travelling Salesperson Problem.

- (a) Plot the two RTD graphs and briefly discuss the differences in performance between the two algorithms seen from the RTD plots.
- (b) Report the median run-time for both RTDs and analyse the statistical significance of the observed values using an appropriate statistical test. Briefly explain your solution.
- (c) Apply an appropriate statistical test to investigate whether the two empirical RTDs are different (the answer is obvious from part (a), but here you should verify this formally using the correct test). Briefly explain your solution.

**Problem 5 (10 marks)**

What is a ROC curve and how can it be used for the empirical analysis of decision algorithms with error? Please give a precise and concise (brief) answer in your own words. If you like, you can illustrate your answer with drawings (but you don't have to do that). Your answer should be no longer than 1/4–1/2 page.