T-61.5060 EXERCISE 7/2005

In T3 on 30 November 2005 at 12 o'clock.

Recall the notation for the sets of variables: $ABC = \{A, B, C\}$ etc.

1. Investigate the data set "smallilmo" to see if you can find any fragments of order of length 3 from it. Are there any longer fragments? Can you see any patterns?

2. Given a 0-1 data set with n rows and where the attributes are independent and 1 with probability p, what is the expected number of violations of a fragment $A < B < C$.

3. We considered in the lectures the greedy approximation algorithm for the set cover problem. What other types or additional heuristics could one use for this task?