Universitat Politècnica de Catalunya
Facultat d'Informàtica de Barcelona

Degree: Grau en Enginyeria Informàtica **Academic year:** Q1 2022–2023 (Mid-term Exam)
Course: Randomized Algorithms (RA-MIRI) **Date:** October 28th, 2022
Time: 1h 45m

1. **(2.5 points)** Iris recognition is based in measures of the iris image, which is ultimately encoded as a bitvector of 2048 bits. Careful studies have shown that 266 of these are uncorrelated. Thus we could assign a 267-bit code[1] to every existing person for unique identification if the probability that the iris codes of two different people have a large coincidence is extremely low. On the other hand two independent measures of the iris image of a person will produce slightly different iris codes, but the number of differences will be small.

   If we think of iris codes as random bitvectors, two bitvectors drawn at random we expect them to differ in about 133 positions (50% of the 267 bits). Give an upper bound to the probability that two randomly chosen iris codes differ in less than 89 positions (1/3 of the 267 bits).

   N.B. You can leave some results as symbolic constants, e.g., e^{-1} instead of giving the approximation 0.367...

2. **(2.5 points)** Professor Church is about to perform the following experiment with two decks of cards, the "red deck" and the "blue deck", both facing down in a table in front of Prof. Church and shuffled in a way that every permutation of the 52 cards of each deck is equally likely.

   On each round, Professor Church turns over the top card of each deck. If any of the two cards is the three of clubs (3♣) the experiment is over. Otherwise the two cards are put back into their decks, and the decks re-shuffled. What is the expected number of rounds played in this experiment?

   Once the first experiment has been finished, Professor Church picks the 52 cards of the red deck, shuffles them, and starts a second experiment. Again, the experiment will be over when the card on top is the three of clubs. However, if the card on top is not 3♣ then that card is put aside and a new round started. What is the expected number of rounds played in the second experiment?

[1] To round up.

3. **(2.5 points)** Suppose now we set up a third experiment, very much like those of the previous question, with two decks as in the first experiment, but when the two cards on top are $\neq 3\clubsuit$ we put them aside, like in the second experiment. Compute the exact probabilities of the following events:

   (a) The first pair consists of two 3s.

   (b) One card of the first pair is a 3, and the other card is a club ($\clubsuit$).

   (c) The second pair consists of two 3s (given that the first round was not the last!).

   (d) One card of the second pair is a 3, and the other card is a club ($\clubsuit$), given that the first round was not the last.

4. **(2.5 points)** We have been given a box with n bolts and n nuts of different sizes. Each nut matches exactly one bolt and vice versa. All bolts and nuts are almost of the same size, so we cannot tell if one bolt is bigger than another or if one nut is bigger than another, and we don't have any instrument to measure them precisely. However, we can try to match one nut with one bolt, and the outcome is always clear: the nut is either too big, too small or exactly right for the bolt.

   The obvious algorithm —for each nut test every bolt until we find its right match— is too costly: we would need $\Theta(n^2)$ tests in the worst case to match every nut with its corresponding bolt.

   Devise a randomized algorithm that efficiently solves this problem on average. The algorithm must always solve the problem correctly, that is, it never leaves unmatched bolts or nuts. Explain your algorithm with enough detail to justify its correctness and to analyze is expected cost. Show that the expected number of tests for your algorithm satisfies a recurrence of the form

$$T_n = c \cdot n + \sum_{k=1}^{n} \pi_{n,k} \cdot (T_{k-1} + T_{n-k}),$$

   for some constant c and probabilities $\pi_{n,k}$ to be determined. Solve the recurrence using the continuous master theorem.

   Useful formula:

$$\int z \ln(z)\, dz = \frac{z^2 \ln(z)}{2} - \frac{z^2}{4} + C$$