Karlsruhe Institute of Technology

Stefan Walzer, Stefan Hermann
ITI Prof. Sanders
winter term 2025/2026

Exercise Sheet 15 – Peeling

Probability and Computing

Exercise 1 – Deterministic peeling in linear time

(a) The algorithm constructByPeeling is nondeterministic (in the choice of i in the while-loop). Show that nevertheless, for any two possible executions of constructByPeeling, the same set of keys remains unplaced. (In particular, the nondeterminism does not affect success/failure.)

(b) Refine the pseudocode so that a deterministic algorithm results, which clearly has running time $O(n)$. Use suitable auxiliary data structures of your choice.

Exercise 2 – Peeling with 2 hash functions

Suppose we use the peeling algorithm constructByPeeling in a setting with only two hash functions h_1 and h_2. For simplicity we assume that $h_1(x) \neq h_2(x)$ for all $x \in D$, and under this restriction the pairs $(h_1(x), h_2(x))$ are uniformly random and independent for different $x \in D$.[1] Show:

(a) All keys are placed if and only if the following graph is acyclic:

$$G = ([m], \{\{h_0(x), h_1(x)\} \mid x \in S\})$$

Note: G is to be understood as a multigraph.

(b) Let $\alpha > 0$ be a constant and $n = \lfloor \alpha m \rfloor$. Show that (for $m \to \infty$) there is a cycle with probability $\Omega(1)$.
Hint: It suffices to look for cycles of length 2 (i.e., double edges).
See `https://en.wikipedia.org/wiki/Birthdayproblem#Arbitrarynumberofdays`.

Hence there is no peelability threshold as there is for $k \geq 3$, since for no $\alpha = \Theta(1)$ does constructByPeeling succeed with high probability.

[1] To ensure this we can, for example, define $h_2(x) := (h_1(x) + h_{\text{diff}}(x)) \bmod m$ for some $h_{\text{diff}} : D \to [m-1]$.

Exercise 3 – The peeling algorithm does not get stuck only late[2]

For a set of keys $S \subseteq D$ of size $n = |S|$ and hash functions $h_1, h_2, h_3 \sim \mathcal{U}([m]^D)$ we consider the cuckoo graph as in the lecture:

$$G = G_{S,h_1,h_2,h_3} = (S, [m], \{(x, h_i(x)) \mid x \in S, i \in [3]\})$$

We assume only $\alpha = \frac{n}{m} \leq 1$. Let $S' \subseteq S$ be the set of keys that cannot be removed by the peeling algorithm (we have $|S'| \in \{0, \ldots, n\}$). We want to show that $\Pr[|S'| \in \{1, \ldots, \delta m\}] = O(1/m)$ for a constant $\delta > 0$ (to be chosen later).

Intuition: Either $S' = \emptyset$ or $|S'|$ is $\Omega(m)$; everything in between is unlikely.

(a) Observe: $|S'| = 1$ is not possible.

(b) Show: $|N(S')| \leq \frac{3}{2}|S'|$. Here $N(S')$ is the set of all neighbors of S' in G.

(c) Show that there exists a constant C such that for $s \in \{2, \ldots, n\}$ the following holds:

$$p_s := \Pr[\exists X \subseteq S, |X| = s : \exists Y \subseteq [m], |Y| = \lfloor \tfrac{3}{2}s \rfloor : N(X) \subseteq Y] \leq \left(C \cdot \frac{s}{m}\right)^{s/2}.$$

Hint: Simply apply a brute-force union bound over all choices of X and Y. The bound $\binom{n}{k} \leq (\frac{ne}{k})^k$ for binomial coefficients is also useful. Ignore the floor functions; everyone does that.

(d) Show (i) $p_2 + p_3 + p_4 + p_5 = O(1/m)$, (ii) $\sum_{s=6}^{\sqrt{m}} p_s = O(1/m)$, (iii) $\sum_{s=\sqrt{m}}^{m/(2C)} p_s = O(1/m)$.

(e) Choose $\delta = 1/2C$ and show $\Pr[|S'| \in \{1, \ldots, \delta m\}] \leq \sum_{s=2}^{\delta m} p_s = O(1/m)$.

[2] This exercise is somewhat involved. It is mainly on the sheet because otherwise the proof from the lecture would be incomplete.