A language-theoretic approach to the heapability of signed permutations

Gabriel Istrate

gabriel.istrate@unibuc.ro

UNIVERSITY OF BUCHAREST
VIRTUTE ET SAPIENTIA

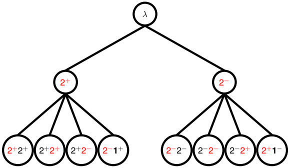
Plan of the presentation

- What exactly do we prove (**only the stringology part**).
- (long detour) Why do we care:

> Sequence-theoretic (open) problem
> (variation of one in CPM'2015 paper) ->
> reduction to string-theoretic prob. ->
> (eventually): experimental study.

- Presentation: **Ideas, not proofs.**
- What's next.

What we study

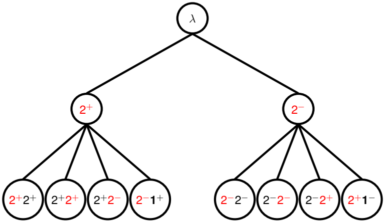


- Signed Hammersley process: $L(H_k^{sign}) \subseteq \Gamma_k^*, \Gamma_k := \{0^+, 0^-, \ldots, k^+, k^-\}$.

- Given current word w:
  1. insert a new letter $z \in \{k^+, k^-\}$ into w.
     Positive/negative particle with k lives.
  2. $z = k^+$ "takes" one life from the closest negative particle p^-, $p \geq 1$, to its right (if any). $z = k^-$: analogous.

- What words generated? Given w, its multiplicity? How does a "typical" word look like? Without signs. G.I., Bonchiş, Rochian, MCU'2018.

Results in the paper in a nutshell



- explicit characterization of words generated by the process: intersection of two DCFL (probably **not** CFL).
- Algorithm for computing multiplicity of a word
- Clarify the connection to our motivating problem.

What we study/prove (2)

Definition: $z \in \Gamma_k^*$ is called k-dominant iff it starts with a letter from the set $\{k^+, k^-\}$, and satisfies:

$$|z|_{k^+} - \sum_{i=1}^{k} i \cdot |z|_{(k-i)^-} + \sum_{i=0}^{k-1} |z|_{i^+} \geq 0 \tag{1}$$

and

$$|z|_{k^-} - \sum_{i=1}^{k} i \cdot |z|_{(k-i)^+} + \sum_{i=0}^{k-1} |z|_{i^-} \geq 0 \tag{2}$$

at least one of the inequalities being strict, namely the one that corresponds to the first letter of z.

Characterization of words

THEOREM: word $z \in \Gamma_k^*$ is generated by the signed Hammersley process if and only if z and all its nonempty prefixes are k-dominant

Corollary: For $k \geq 1$, if $L(H_k^{sign})$ is the language of generated words, there exist two deterministic context-free languages (in fact L_1, L_2 are even deterministic one-counter languages, (Valiant, 1975)) s.t. $L(H_k^{sign}) = L_1 \cap L_2$.

Multiplicity of words

- Paper: **Algorithm based on dynamic programming.**
- We "reverse" possible derivations: find all the preimages of a given word in the signed Hammersley process.
- conceptually simple, tedious.
- In the paper, not in this presentation !

Why do we care

- The (classical) Ulam-Hammersley problem.
- Heapability, and the Ulam-Hammersley problem for heapable sequences

Results in this paper: motivated by the Ulam-Hammersley problem for **signed** heapable sequences

- (CPM'2015) The golden ratio conjecture and a "physics-like" argument for it.
- MCU'2018: Attempt to prove this conjecture via formal power series. Made (baby) first-steps.
- This paper: Start similar program for signed permutations.

Starting Point

Longest Increasing Subsequence

3 2 5 7 1 6 9

Patience sorting.

Another (greedy, also first-year) algorithm:

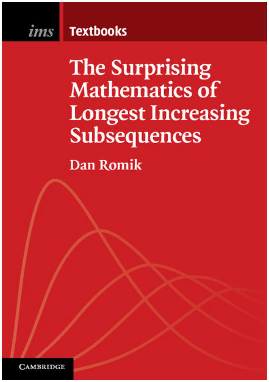
> Start (greedily) building decreasing piles. When not possible, start new pile.

Size of LIS = # of piles in patience sorting.

The Ulam-Hammersley problem (for random permutations)

What is the LIS of a random permutation ?

$$E_{\pi \in S_n}[LIS(\pi)] = 2\sqrt{n} \cdot (1 + o(1)).$$



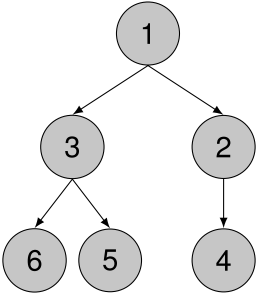
- Logan-Shepp (1977), Veršik-Kerov (1977), Aldous-Diaconis (1995)
- Very rich problem. Connections with nonequilibrium statistical physics and Young tableaux

From (increasing) sequences to heaps

Byers, Heeringa, Mitzenmacher, Zervas (ANALCO'2011)

Sequence of integers A is heapable if it can be inserted into binary heap-ordered tree (not necessarily complete), always as leaf nodes.

Example: 1 3 2 6 5 4   Counterexample: 5 1 . . .



The Ulam-Hammersley problem for heapable sequences

- (Dilworth, patience sorting): $LIS(\pi) =$ minimum number of decreasing piles in a partition of π.

$HEAPS_k =$ minimum number of k-heaps in a partition of π

Ulam-Hammersley problem for heapable sequences:

What is the scaling of $E_{\pi \in S_n}[HEAPS_k(\pi)]$, $k \geq 2$?

A beautiful golden-ratio conjecture (CPM'2015)

For $k \geq 2$ there exists $\lambda_k > 0$ such that

$$\lim_{n \to \infty} \frac{E[HEAPS_k(\pi)]}{ln(n)} = \lambda_k$$

Moreover

$$\lambda_2 = \frac{1+\sqrt{5}}{2}$$

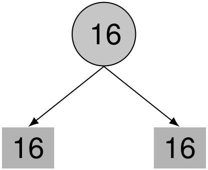
is the golden ratio.

Status of the conjecture

- "Physics-like" nonrigorous argument, includes prediction for value of constant λ_k.
- Basdevant et al. (2016, 2017) rigorously establishes logarithmic scaling, but not the value of the constant.
- (MCU'2018) Language-theoretic perspective (that we generalize in this paper). Experimentally supports golden-ratio value of λ_2.

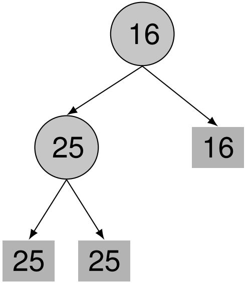
Connection to Hammersley's process: Patience heaping

16, 25, 18, 2, 4, 35, 3, 7, 32, 9, 20



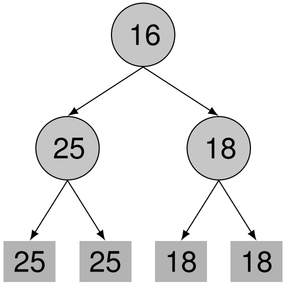
Connection to Hammersley's process: Patience heaping

16, 25, 18, 2, 4, 35, 3, 7, 32, 9, 20



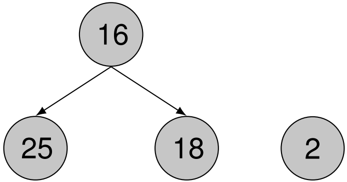
Connection to Hammersley's process: Patience heaping

16, 25, 18, 2, 4, 35, 3, 7, 32, 9, 20



Connection to Hammersley's process: Patience heaping

16, 25, 18, 2, 4, 35, 3, 7, 32, 9, 20

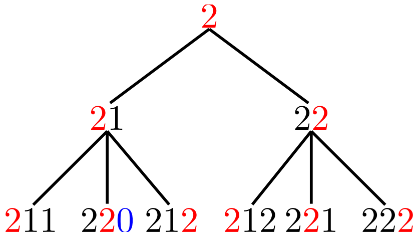


- 2 does not kill any slots!
- the number of heaps increases by 1.

Hammersley's process:

- Particles: slots in patience heaping
- Choose a random position. Put there a 2. Remove 1 from the closest nonzero digit to the right (if any).

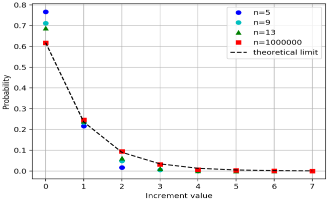
$$E[\Delta(\#\text{ of heaps})] = \frac{1+E[\#\text{ of trailing zeros of } w]}{n+1}$$



A "physicist's explanation" for the golden-ratio conjecture

- $W_n =$ (prefix+) random word in the process at time n.
- $n \to \infty$: *Limit of* W_n = compound Poisson process. $W_n =$ random string of 0,1,2 (densities c_0, c_1, c_2).
- *Assuming this:* Distribution of # of trailing zeros: asymptotically geometric

If I can sample **exactly** from the distribution of *w* then I can compute the scaling constant!



- From this: $E[\Delta(\#\ heaps.)\ at\ stage\ n] \sim \frac{1+\sqrt{5}}{2} \cdot \frac{1}{(n+1)}$.

This paper: heapability of signed permutations

A signed permutation of order n is a pair (σ, τ), with $\sigma \in S_n$ and $\tau : [n] \to \{\pm 1\}$ being a *sign function*.

Given integer $k \geq 1$, signed permutation (σ, τ) is called $\leq$ *k-heapable* if one can successively insert elements of (σ, τ) into k (min) heap-ordered binary trees (not necessarily complete) $H_0, H_1, \ldots, H_{k-1}$ such that within each heap signs alternate between parent and child nodes

Random model S_n^p for signed permutations:

- Generate random $\sigma \in S_n$.
- Constant $p \in [0, 1]$.
- Each τ_i: 1 independently w.p. p (-1 with probability $1 - p$).

Question: expected number of heaps in a heap decomposition of a random signed permutation $(\sigma, \tau) \in S_n^p$?

This paper

- Proved (Theorem 4, paper) that a version of patience heaping is exact for the Ulam-Hammersley problem for signed permutations
- Consequently, **the signed Hammersley process relevant for the problem we investigate.**
- Did not do (yet) the experimental studies ...
- .... but the multiplicity algorithm from the paper good (in principle) for them.

Conclusions/What's next:

Rich problem with many open questions

- Next: experimental studies !
- Problem studied here **not** variant of unsigned one.
- Still logarithmic scaling may still be possible ...
- ... but not at endpoints ($p = 0, p = 1$)! $HAM_k(\pi) = n$.

Thank you. Questions ?

gabriel.istrate@unibuc.ro