On a greedy algorithm for non-deterministic walks with several letters

Sergey Dovgal, Philippe Duchon, Mohamed Lamine Lamali

LaBRI, Université de Bordeaux

Lattice Path Conference 21–25 June 2021. Presentation times for the poster: Wednesday and Thursday 1h30–2h30 pm

Abstract

A one-dimensional **N-walk with several letters** is a simplified version of a non-deterministic push-down automaton whose underlying digraph of transitions is a path. Each of the non-deterministic steps is a set of the possible stack operations. We study the **acceptance probability of an empty word** in the model where non-deterministic steps are chosen with given probabilities. We prove that the probability that the greedy algorithm finds an excursion compatible with a given N-path undergoes a coarse **phase transition**.

Non-deterministic walks with several letters

- **An N-walk is a sequence of sets of admissible steps.**
- A **step** of Dyck type can be of the form $\uparrow_{\mathbf{x}}$ (adding a letter) or $\downarrow_{\mathbf{x}}$ (removing a letter), where $\mathbf{x} \in \mathbf{\Sigma}$. Furthermore, a step can be of the form $\uparrow_{\mathbf{x}|\mathbf{y}}$: add a letter $\mathbf{x}$ if the stack head is $\mathbf{y}$.
- An **excursion with several letters** is a sequence of stack states, where each step either removes a top stack letter, or adds one.
- An **N-excursion** is an N-walk compatible with at least one deterministic excursion.

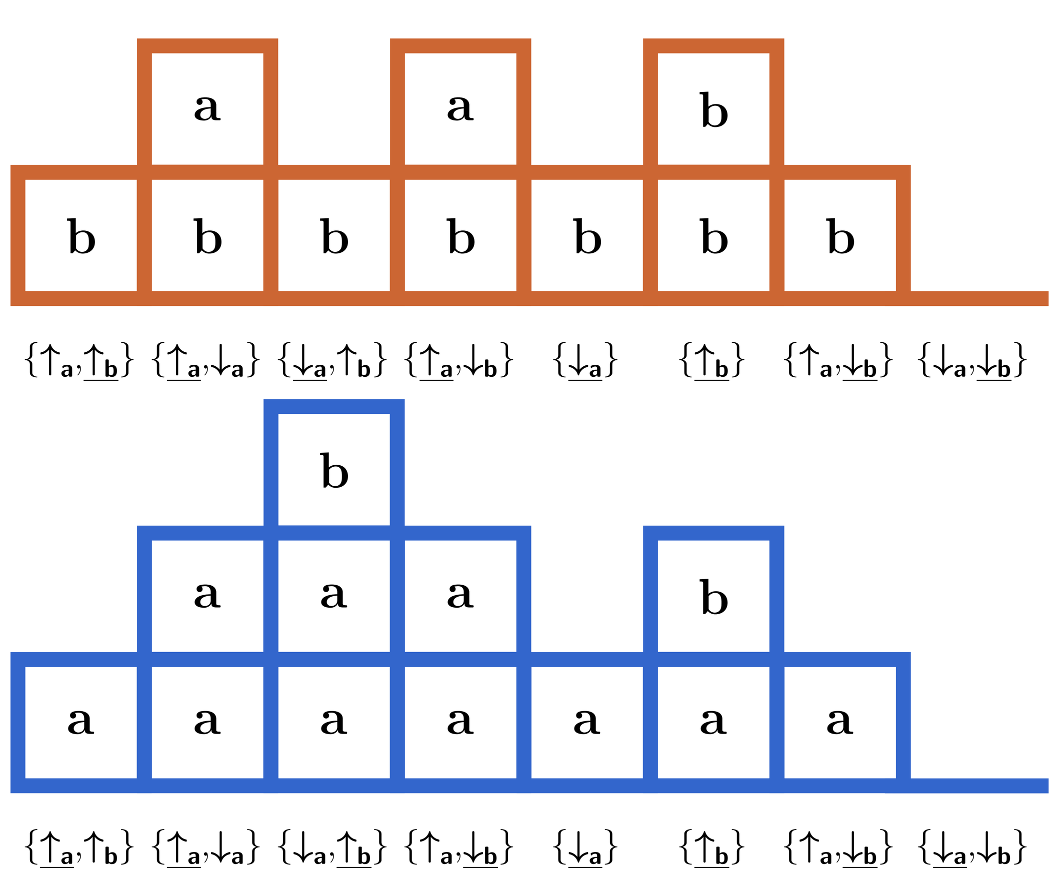

Figure 1: Two realisations of a non-deterministic excursion

Choice of the step set

Consider an admissible N-step set $\mathcal{S} \subset 2^{\cup_{\mathbf{x}\in\Sigma}\{\uparrow_{\mathbf{x}},\downarrow_{\mathbf{x}}\}}$.
In the weighted model, each N-step is chosen independently from some distribution over $\mathcal{S}$.
We are looking for distributions $\big(\mathbb{P}(s)\big)_{s\in\mathcal{S}}$ that yield a *positive limit probability for an N-excursion.*

1. $\mathbb{P}(\{\uparrow_{\mathbf{x}}\})\mathbb{P}(\{\downarrow_{\mathbf{y}}\}) = 0$ for any distinct $\mathbf{x}, \mathbf{y} \in \mathbf{\Sigma}$;
2. $\mathbb{P}(\{\uparrow_{\mathbf{x}}\})\mathbb{P}(\{\uparrow_{\mathbf{y}}\downarrow_{\mathbf{y}}\})\mathbb{P}(\{\downarrow_{\mathbf{x}}\}) = 0$ (idem)
3. Any combination allowing to *get stuck* in a finite number of moves is forbidden (except when the stack is empty).

Running example. N-step set with *two letters*:

$$\mathcal{S} = \Big\{\{\uparrow_{\mathbf{a}}\},\{\uparrow_{\mathbf{b}}\},\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\},\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\},\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}\Big\}$$

Example from left-right mirror symmetry.

$$\hat{\mathcal{S}} = \Big\{\{\downarrow_{\mathbf{a}}\},\{\downarrow_{\mathbf{b}}\},\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\},\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\},\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}\Big\}$$

Greedy algorithm

A greedy algorithm always tries to *remove a letter* whenever possible. **Action probabilities** for the greedy algorithm, conditioned on the top letter of the stack (for the running example):

$$\begin{aligned}
&\mathbb{P}(\uparrow_{\mathbf{a}} \,|\, \mathbf{a}) = \mathbb{P}(\{\uparrow_{\mathbf{a}}\})\\
&\mathbb{P}(\uparrow_{\mathbf{b}} \,|\, \mathbf{a}) = \mathbb{P}(\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{b}}\})\\
&\mathbb{P}(\downarrow_{\mathbf{a}} \,|\, \mathbf{a}) = \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\})\\
&\mathbb{P}(\uparrow_{\mathbf{a}} \,|\, \mathbf{b}) = \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{a}}\})\\
&\mathbb{P}(\uparrow_{\mathbf{b}} \,|\, \mathbf{b}) = \mathbb{P}(\{\uparrow_{\mathbf{b}}\})\\
&\mathbb{P}(\downarrow_{\mathbf{b}} \,|\, \mathbf{b}) = \mathbb{P}(\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\})
\end{aligned}$$

- The *drift* is the expected height of the final stack state divided by the length of the walk.
- Greedy algorithm fails if the drift is positive
- LR $\wedge$ RL drift conditions provide the correct phase transition threshold when $|\Sigma| = 1$.

Central result (presented for the running example)

$$\text{Let } \Delta = \det\begin{pmatrix} \mathbb{P}(\uparrow_{\mathbf{a}} \,|\, \mathbf{a}) - \mathbb{P}(\downarrow_{\mathbf{a}} \,|\, \mathbf{a}) & \mathbb{P}(\uparrow_{\mathbf{b}} \,|\, \mathbf{a}) \\ \mathbb{P}(\uparrow_{\mathbf{a}} \,|\, \mathbf{b}) & \mathbb{P}(\uparrow_{\mathbf{b}} \,|\, \mathbf{b}) - \mathbb{P}(\downarrow_{\mathbf{b}} \,|\, \mathbf{b}) \end{pmatrix}.$$

If the greedy algorithm finds a compatible excursion with positive limit probability, then $\Delta > 0$.
The boundary between the regions where the limit probability is positive and zero satisfies $\Delta = 0$.

Proof idea: recurrences and generating functions (GFs)

Let $a_n, b_n, w_n^{\mathbf{a}}, w_n^{\mathbf{b}}, u_n^{\mathbf{a}}, u_n^{\mathbf{b}}$ denote the probabilities that the top letter of a greedy walk of size n is $\mathbf{a}$ or $\mathbf{b}$, and the weight of strictly positive excursions starting with the letters $\mathbf{a}$ or $\mathbf{b}$, and non-negative excursions starting with the letters $\mathbf{a}$ or $\mathbf{b}$ or length n. Let $A(z), B(z), W_{\mathbf{a}}(z), W_{\mathbf{b}}(z), U_{\mathbf{a}}(z)$ and $U_{\mathbf{b}}(z)$ be their GFs. Then,

$$\begin{aligned}
A(z) &= \frac{z}{1-z}\big(\mathbb{P}(\{\uparrow_{\mathbf{a}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\})\big) - zA(z)\mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}}\})\\
&\quad - z(A(z)+B(z))\mathbb{P}(\{\uparrow_{\mathbf{a}},\downarrow_{\mathbf{a}},\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}) + A(z)W_{\mathbf{a}}(z),\\
B(z) &= \frac{z}{1-z}\big(\mathbb{P}(\{\uparrow_{\mathbf{b}}\}) + \mathbb{P}(\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\})\big) - zB(z)\mathbb{P}(\{\uparrow_{\mathbf{b}},\downarrow_{\mathbf{b}}\}) + B(z)W_{\mathbf{b}}(z),\\
W_{\mathbf{x}}(z) &= z^2 \sum_{\mathbf{y}\in\mathbf{\Sigma}} U_{\mathbf{y}}(z)\mathbb{P}(\uparrow_{\mathbf{y}} \,|\, \mathbf{x})\mathbb{P}(\downarrow_{\mathbf{y}} \,|\, \mathbf{y}) \quad \text{for} \quad \mathbf{x} \in \mathbf{\Sigma},\\
U_{\mathbf{x}}(z) &= 1 + z^2 U_{\mathbf{x}}(z) \sum_{\mathbf{y}\in\mathbf{\Sigma}} U_{\mathbf{y}}(z)\mathbb{P}(\uparrow_{\mathbf{y}} \,|\, \mathbf{x})\mathbb{P}(\downarrow_{\mathbf{y}} \,|\, \mathbf{y}) \quad \text{for} \quad \mathbf{x} \in \mathbf{\Sigma}. \qquad (*)
\end{aligned}$$

$U_{\mathbf{x}}(z)$ satisfy a system of quadratic equations. Next, $W_{\mathbf{x}}(\pm 1) = 1 - \mathbb{P}(\downarrow_{\mathbf{x}} \,|\, \mathbf{x})$ are *valid formal solutions*. If they correspond to *principal branches* of the generating functions, then the greedy algorithm succeeds with a positive probability. That is not always the case: another branch will become principal after coalescence of the *multiple roots*, which happens when $\Delta = 0$.

Conclusions, open problems

Existence of an excursion compatible with an N-walk is a particular case of the Constraint Satisfiability Problem (CSP). Study of the phase transitions in random CSP is a rich ongoing research topic.

1. According to the simulations for $\mathcal{S}$ from the running example, the condition $\Delta > 0$ is necessary and sufficient for the limit **true probability that an N-walk is an N-excursion** to be strictly positive.
2. Does the phase transition exist for the true probability of acceptance when $|\Sigma| \geq 2$?
3. If so, is the true phase transition threshold described by a conjunction of the Left-Right and Right-Left greedy algorithm thresholds when $|\Sigma| \geq 2$?
4. If so, can we describe the asymptotics of the true acceptance probability in the *subcritical case* (when it tends to zero) and in the *supercritical case* (when it is positive)?
5. If n is the length of an N-walk, and, for example, $\mathbb{P}(\{\uparrow_{\mathbf{x}}\}) = \Theta(n^{-1})$ together with $\mathbb{P}(\{\downarrow_{\mathbf{y}}\}) = \Theta(n^{-1})$, how does the phase transition threshold change?

References

[1] Élie de Panafieu, Mohamed Lamine Lamali, and Michael Wallner.
Combinatorics of nondeterministic walks of the Dyck and Motzkin type.
In *ANALCO*, pages 1–12, 2019.

Acknowledgements

Many thanks to Élie de Panafieu. The 1st and the 3rd authors were supported by the HÉRA project, funded by The French National Research Agency. Grant no.: ANR-18-CE25-0002.

LABORATOIRE BORDELAIS DE RECHERCHE EN INFORMATIQUE LaBRI

université de BORDEAUX