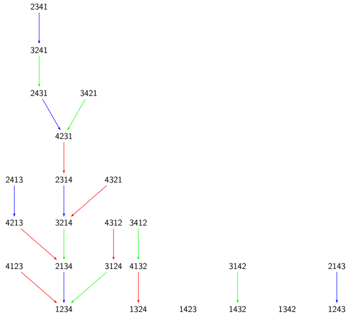
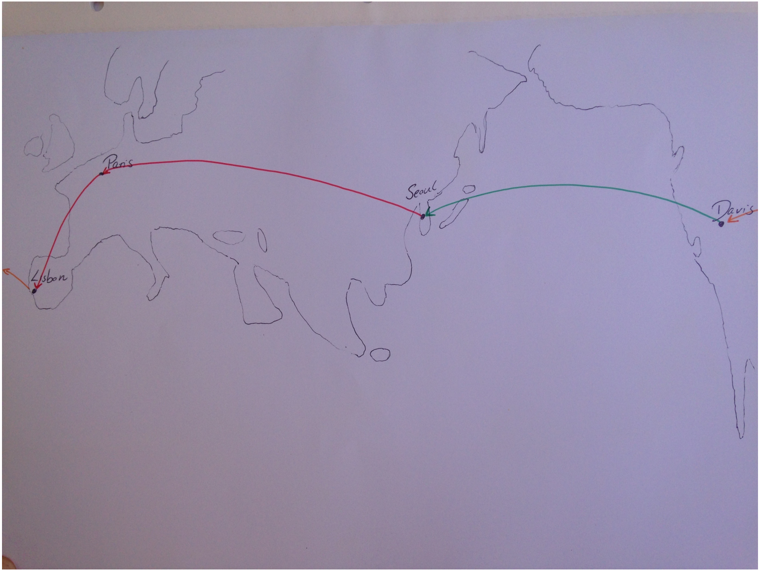
A Markov chain on semaphore codes and the fixed point forest

A. Schilling, UC Davis

International Conference on Semigroups and Automata 2016
Celebrating the 60th birthday of Gracinda Gomes and Jorge Almeida
June 24, 2016



Map



Outline

1. *A Markov chain on semaphore codes*
   in joint work with John Rhodes and Pedro Silva
   arXiv:1509.03383 and arXiv:1604.00959, to appear in IJAC

2. *The fixed point forest*
   in joint work with Tobias Johnson and Erik Slivken
   arXiv:1605.09777 submitted

Appearance of *probability* and *combinatorics* in semigroup theory!

de Bruijn graph

A finite alphabet

de Bruijn graph:

vertices words in A of length k
edge $a_1 \cdots a_k \xrightarrow{a} a_2 \cdots a_k a$

random walk:

$v \xrightarrow{a} w$ with probability $\pi(a)$

transition matrix:

$\mathcal{T}_{v,w} = \pi(a)$ if $v \xrightarrow{a} w$

Stationary distribution $I\mathcal{T} = I$?
Answer: $I = (\prod_{a \in w} \pi(a))_{w \in A^k}$

Action of semigroup

Semigroup: $F(A,k) = A^1 \cup A^2 \cup \cdots \cup A^k = A^{\leqslant k}$
with product taking last k letters of concatenation

Action: $F(A,k)$ acts on A^k as

$$a_1 \ldots a_k \cdot a = a_2 \ldots a_k a \qquad \text{for } a \in A$$

Resets: elements in semigroup that act as constant maps

Here A^k

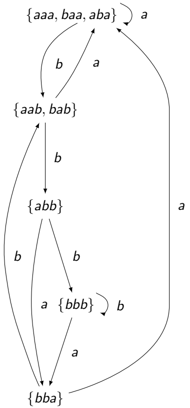
Right congruences

Motivation: Capture information that matters!

Example

$A = \{a, b\}$

$\mathrm{RC}(A^3) = \{\{aaa, baa, aba\}, \{bba\}, \{aab, bab\}, \{abb\}, \{bbb\}\}$



Right congruences

$$\mathrm{RC}(A^3) = \{\{aaa, baa, aba\}, \{bba\}, \{aab, bab\}, \{abb\}, \{bbb\}\}$$

Transition matrix:

$$\mathcal{T} = \begin{pmatrix} \pi(a) & 0 & \pi(b) & 0 & 0 \\ \pi(a) & 0 & \pi(b) & 0 & 0 \\ \pi(a) & 0 & 0 & \pi(b) & 0 \\ 0 & \pi(a) & 0 & 0 & \pi(b) \\ 0 & \pi(a) & 0 & 0 & \pi(b) \end{pmatrix}$$

Stationary distribution by lumping:

$$I = (\pi(a)^2 + \pi(a)^2\pi(b), \pi(a)\pi(b)^2, \pi(a)\pi(b), \pi(a)\pi(b)^2, \pi(b)^3)$$

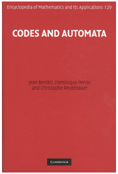
Goal: hitting time

Approach

- right congruences form a *lattice* under inclusion (meets and joins exist)
- approximation by *special congruences*
- *special congruences* $\longleftrightarrow$ semaphore codes

Suffix codes

see *Berstel, Perrin, Reutenauer* Codes and Automata



A finite alphabet
A^+ free semigroup with generators in A
A^* free monoid with generators in A

Definition

u suffix of v $\quad\Leftrightarrow\quad$ $\exists w \in A^*$ such that $wu = v$

Definition

Suffix code $\mathcal{C}$ is subset $\mathcal{C} \subseteq A^+$ such that elements in $\mathcal{C}$ are pairwise incomparable in suffix order (antichain)

Semaphore codes

Definition

A *semaphore code* is a suffix code $\mathcal{S}$ over A that has a right action:

$$u \in \mathcal{S}, a \in A \qquad \Rightarrow \qquad ua \quad \text{has suffix in } \mathcal{S}$$

Example

$\mathcal{S} = \{ba^j \mid j \geqslant 0\} = ba^*$

$ba^j \cdot a = ba^{j+1}$
$ba^j \cdot b = b$

Codes and ideals

Definition

$\mathcal{L} \subseteq A^+$ is a left ideal if $u\mathcal{L} \subseteq \mathcal{L} \; \forall u \in A^*$

suffix code = suffix minimal elements of left ideal

Definition

$\mathcal{I} \subseteq A^+$ is a ideal if $u\mathcal{I}v \subseteq \mathcal{I} \; \forall u, v \in A^*$

Connection to *semaphore codes*:
Take $u = a_j \cdots a_1 \in \mathcal{I}$. Find unique index $1 \leqslant i \leqslant j$ such that

$$a_{i-1} \cdots a_1 \notin \mathcal{I} \quad \text{but} \quad a_i \cdots a_1 \in \mathcal{I}$$

Then $a_i \cdots a_1$ is a code word.

$\mathrm{Sem}(A^k)$ set of semaphore codes with ideal in $A^{\leqslant k}$

Semaphore codes and right congruences

$u, v \in A^k$: $u \sim_{\mathcal{S}} v$ if u and v have a common suffix in $\mathcal{S}$

$\sim_{\mathcal{S}}$ defines a right congruence on A^k

Example

$A = \{a, b\}$

$$\mathcal{S} = \{aa, ab, aba, bba, abb, bbb\} \qquad \textit{semaphore code}$$

$\mathcal{S}$ yields right congruence in $\mathrm{RC}(A^3)$:

$$\{aaa, baa\}, \{aab, bab\}, \{aba\}, \{bba\}, \{abb\}, \{bbb\}$$

All congruences resulting from semaphore codes are called special right congruences $\mathrm{SRC}(A^k)$.

Approximation

$\mathrm{RC}(A^k)$ set of right congruences
$\mathrm{SRC}(A^k)$ set of special right congruences

$\mathrm{SRC}(A^k)$ full sublattice (top and bottom agree) of $\mathrm{RC}(A^k)$

Each $\rho \in \mathrm{RC}(A^k)$ has a unique largest lower (finer) approximation $\underline{\rho} \in \mathrm{SRC}(A^k)$

$$\underline{\rho} = \bigvee_{\substack{\tau \in \mathrm{SRC}(A^k) \\ \tau \subseteq \rho}} \tau \qquad \text{(join)}$$

Example

$\rho = \{\{aaa, baa, aba\}, \{bba\}, \{aab, bab\}, \{abb\}, \{bbb\}\} \in \mathrm{RC}(A^3)$
Approximation:
$\underline{\rho} = \{\{aaa, baa\}, \{aba\}, \{bba\}, \{aab, bab\}, \{abb\}, \{bbb\}\}$

Random walk on semaphore codes

probability distribution: $\pi \colon A \to [0,1]$

transition matrix: $\mathcal{T} = \sum_{a\in A} \pi(a)\mathcal{T}(a)$
with $\mathcal{T}(a)_{s,s\cdot a} = 1$ and 0 else

Theorem (RSS 2015)

Probability that word of length ℓ is reset:

$$P(\ell) = \sum_{\substack{s\in\mathcal{S}\\ \ell(s)\leqslant\ell}} \pi(s)$$

Observation

ρ and approximation $\underline{\rho}$ have same hitting time!

Random walk on semaphore codes

Example

semaphore code: $\mathcal{S} = ba^*$
resets: all words w unless $w = a^\ell$
probability that word of length 3 is reset:

$$P(3) = \pi(b) + \pi(b)\pi(a) + \pi(b)\pi(a)^2 = 1 - \pi(a)^3$$

Stationary distribution

Theorem (RSS 2015)

Stationary distribution

$$I = (\pi(s))_{s \in \mathcal{S}}$$

Transition matrix not diagonalizable

Example

$\mathcal{S} = \{a, ab, abb, bbb\}$
Jordan form

$$\left[\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

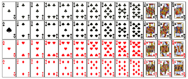
Further work

- Semaphore codes attached to *Turing machines*
- *Profinite limits*
- Characterization of polynomial time Turing machines in this framework, including the natural semaphore codes action
- *P versus NP*??

Outline

2. *The fixed point forest*
   in joint work with Tobias Johnson and Erik Slivken
   arXiv:1605.09777 submitted

Partial Sorting Algorithm

Deck of n cards labelled $\{1, 2, \ldots, n\}$ 

Take top card and move it to slot of its value

Example

$3142 \longrightarrow 1432$

In general, view deck of cards as a permutation

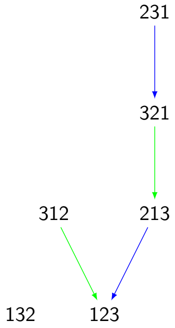
$$\pi(1)\pi(2)\ldots\pi(n) \in \mathfrak{S}_n \quad \longrightarrow \quad \pi(2)\ldots\pi(1)\ldots\pi(n)$$

Fixed Point Forest

- Each permutation eventually *sorted* to permutation with $\pi(1) = 1$.
- Opposite direction: choose *fixed point* and move it to front

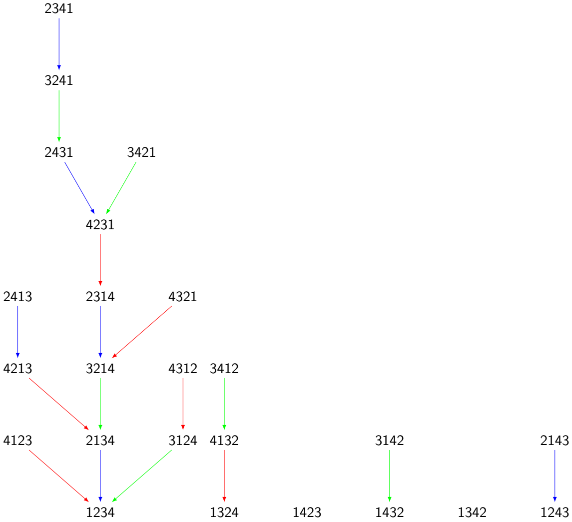
$\rightsquigarrow$ *Fixed point forest F_n* with permutation with $\pi(1) = 1$ as roots and derangements as leaves

Example (Fixed point forest F_3)



Fixed Point Forest

Example (Fixed point forest F_4)



History

Gwen McKinley UC Davis Undergraduate Thesis 2015
started as REU project at Missouri State University by Les Reid
(problem contributed by Gerhardt Hinkle)

Theorem (McKinley 2015)

- *Longest path in F_n of length $2^{n-1} - 1$ starting at $23\ldots n1$*
- *"Fractal structure"*
- *Size of tree containing $12\ldots n$ between $(n-1)!$ and $e(n-1)!$*

Open Problem

Average number of moves to root?

Goal

- Study of *local structure* of tree at random permutation π_n as $n \to \infty$
- *Stein's method*: weak convergence to tree of independent Poisson processes
- *Longest path to leaf*: geometric distribution with mean $e - 1$
- *Shortest path to leaf*: Poisson distribution with mean 1

Moving towards leaves

Recall: choose *fixed point* and move to front

Lemma

Shortest path from π_n to leaf obtained by always bumping rightmost fixed point

Example

$32415 \to 53241 \to 45321 \to 34521$
Shortest path is not unique:
$32415 \to 23415 \to 52341 \to 45231$

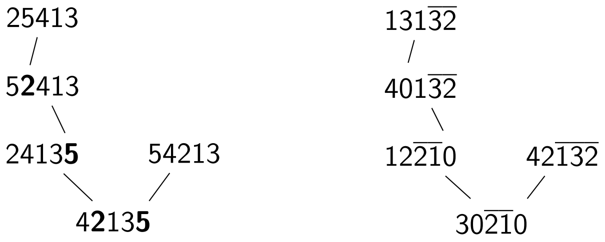
Lemma

Longest path from π_n to leaf obtained by always bumping leftmost fixed point

Remark

Longest path to leaf is unique!

Moving towards leaves in tree $T(\pi)$



Definition

$\pi \in \mathfrak{S}_n$

$\pi(i)$ is *k-separated* if $\pi(i) = i + k$

Structure of $T(\pi)$ up to level ℓ

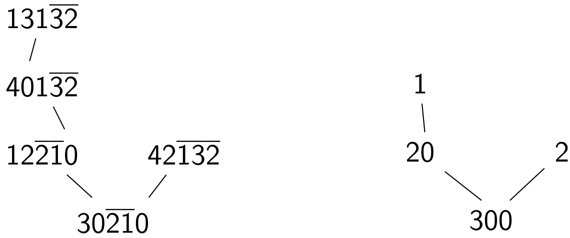
keep track of *k-separated* points for $0 \leqslant k \leqslant \ell$

or *words* in letters $0, 1, \ldots, \ell$

Limiting tree

Algorithm

- *pick* a 0 and remove
- *decrease* all letters to left of 0 by one



Remark

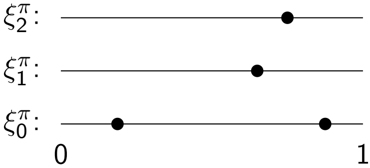
- *This forgets that* 0*-separated points in permutation at position* i *creates* $(i-1)$*-separated point.*
- *This is unlikely in limit* $n \to \infty$.

Poisson point processes

For each k, ξ_k^π represents the *k-separated points* in $[0,1]$ by rescaling by $1/n$.

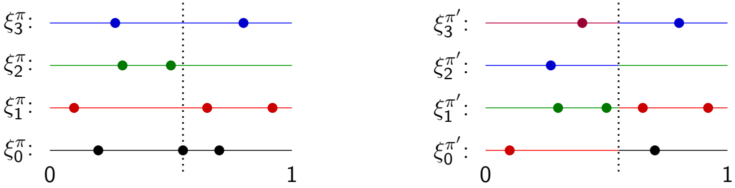
Example

$0120 \longleftrightarrow$



Bumping a fixed point

π is *abstracted permutation*
π' child given by bumping x
$\Rightarrow$ point processes $\xi_k^{\pi'}$ equals ξ_{k+1}^{π} on $[0, x)$ and ξ_k^{π} on $(x, 1]$



Results

T: tree of independent Poisson processes

Theorem (JSS 2016)

F_n weakly converges to T as $n \to \infty$

weak or *Benjamini-Schramm*: k-neighborhood of F_n converges in distribution to k-neighborhood of T

Results

L_n: length of longest path to leaf

Theorem (JSS 2016)

Distribution of L_n converges weakly to geometric distribution G with mean $e - 1$.

$$\mathbf{E}L_n^p \to \mathbf{E}G^p \qquad \forall p > 0$$

M_n: length of shortest path to leaf

Theorem (JSS 2016)

Distribution of M_n converges weakly to Poisson distribution P with mean 1.

$$\mathbf{E}M_n^p \to \mathbf{E}P^p \qquad \forall p > 0$$

Open questions

- T_n tree containing $12\ldots n$ (largest)

$$\frac{1}{n} \leqslant \mathbf{P}[\pi_n \in T_n] \leqslant \frac{e}{n},$$

Limit of $n\mathbf{P}[\pi_n \in T_n]$ as $n \to \infty$

- R_n distance from π_n to the base of its tree in the fixed point forest. *Limiting asymptotics* of $\mathbf{E}R_n$?
- *Random path* from root to leaf. Distribution of the number of steps before reaching a leaf?