NON UNIFORM DISTRIBUTIONS

GENERAL FRAMEWORK

Let S be a set, $\mathbb{P}_n$ be probability laws on S.

DENSITY MODEL: $S = \{$ all finite tuples of (cyclically) reduced words of all lengths and number of entries $\}$

$\mathbb{P}_n$ = uniform probability supported on $\lfloor |CR_n|^d \rfloor$-tuples of elements of CR_n

FEW RELATORS MODEL: fix k, $S = CR^k$ k-tuples of words of whatever length

$\mathbb{P}_n$ = uniform probability supported on $CR^k_{\leq n}$

k-tuples of length $\leq n$

UNIFORM: In both models for each n, two tuples of words of length n (or $\leq n$) have the same probability of being chosen. We will generalize this.

NON UNIFORM DISTRIBUTIONS

For each n let:

$\mathbb{R}_n$ be a probability law on R_n (or CR_n)

Π_n be prob law on the set of tuples of positive integers.

For $\vec{h} \in R_n$ we set

$$\mathbb{P}_n(\vec{h}) = \underbrace{\Pi_n(|h_1|, \dots, |h_k|)}_{\downarrow} \prod_{i=1}^{k} \underbrace{\mathbb{R}_{|h_i|}(h_i)}_{\downarrow}$$

the probability of choosing a tuple with words of length $|h_1| \cdots |h_k|$ among all tuples of natural numbers

↓

Different Π_n's can make $\mathbb{P}_n$ biased towards words of some specified length.

the probability of selecting this word among all the words of its same length.

↓

Different $\mathbb{R}_n$'s can make $\mathbb{P}_n$ more biased towards words of a particular structure

EXAMPLES

•) Few relators: for fixed k:

Π_n = supported on k-tuples and length $\leq n$

and $\Pi_n(\ell_1 \cdots \ell_k) = \prod_{i=1}^{k} \frac{|R_{\ell_i}|}{|R_{\leq n}|}$

$\mathbb{R}_n$ = uniform on R_n (CR_n)

•) <u>Density</u>:

fix $d \in [0,1]$,

Π_n = supported on the $\lfloor (2r-1)^{dn} \rfloor$ - tuple $(n, \dots, n)$

$\mathbb{R}_n$ = uniform on CR_n

<u>Rmk</u>: $\mathbb{P}_n$ just depends on the length of the words, and this in turn depends on how many words of a given length are there in words of length at most n. For a fixed length each word is treated the same.

PREFIX HEAVY DISTRIBUTIONS

Let R be the set of all words. Given $u \in R$ let $P(u) = \{$ words which start with $u \}$.

<u>Def</u>: Let $C > 0$, $\alpha < 1$. A sequence of prob measures $\mathbb{R}_n$, supported on R_n (or CR_n) is (C, α) <u>PREFIX HEAVY</u> if

$$\forall u, v \in R \qquad \mathbb{R}_n (P(uv) \mid P(v)) \leq C \alpha^{|v|}$$

If $\mathbb{R}_n$ are (C, α) prefix heavy, and $\mathbb{P}_n$ defined as above with any Π_n, we also say that $\mathbb{P}_n$ is (C, α)-prefix heavy.

RMK : If $u = \varepsilon$ empty word, the condition is

$$\forall v \in R \quad \mathbb{R}_n(P(v)) \leq C\alpha^{|v|}$$

If we choose $n = |v|$, we get $P(v) \cap R_n = \{v\}$
thus $\mathbb{R}_n(v) \leq C\alpha^n$

And if I take v_n st $|v_n| = n$ then
$\mathbb{R}_n(v_n) \to 0$ exp fast
i.e., probability of picking long words decreases exp fast

$$\mathbb{P}_n(v_n) \leq \pi_n |v_n| \, C\alpha^n \longrightarrow 0.$$

<u>Rmk</u> : Since $|R_n|$ grows exponentially fast in n,
if $\mathbb{R}_n$ are uniform on R_n one can see that they are prefix heavy with $C = 1$ $\alpha = \frac{1}{2r-1}$
(Exact formula for $\mathbb{R}_n(\cdot|\cdot)$ in paper by Bassino et al, [6] in the notes).

<u>Def</u> : If $\vec{h} = (h_1 \cdots h_k)$ is a tuple of reduced words we let

$$\text{size}(\vec{h}) = k \qquad MIN(\vec{h}) = \min_{i=1\dots k} |h_i| \quad MAX(\vec{h}) = \max_{i=1\dots k} |h_i|$$

We see these as random variables on the set of tuples of (cyclically) reduced words.

<u>Def</u> : We say that $\mathbb{R}_n$ (and $\mathbb{P}_n$) does not ignore cyclically reduced words if $\liminf \mathbb{R}_n(C R_n) > 0$

We can state the following which is a generalization for the few relator model.

theorem (Bassino et al. 4.1, [6])
Let $\mathbb{P}_n$ be probability distributions on tuples of reduced words, which are (C, α) prefix heavy with $C \geq 1$, $0 < \alpha < 1$.
Let $0 < \lambda < 1/2$.

-) If $\mathbb{P}_n(\text{SIZE}^2 \cdot \alpha^{\frac{MW}{2}} > y_n) \to 0$ for some sequence $y_n \to 0$, then a random tuple of reduced words generically satisfies the central tree property and freely generates a subgroup of A.

-) If $\mathbb{P}_n(\text{SIZE}^2 \cdot \text{MAX}^2 \alpha^{\lambda MW} > y_n) \to 0$ then a random tuple of reduced words generically generates a malnormal subgroup of $F(A)$

-) If $\mathbb{P}_n$ does not ignore cyclically reduced words and $\mathbb{P}_n(\text{SIZE}^2 \cdot \text{MAX}^2 \alpha^{\lambda MW} > y_n) \to 0$, then a random tuple of cyclically reduced words generically generates a group with $C'(1/6)$

If $\mathbb{P}_n \to 0$ exp. fast, all the results hold exp. generically.

We now introduce a way of defining non uniform distributions and prefix heavy ones.

MARKOVIAN AUTOMATA

Let X be a finite alphabet and let Q be a finite set of STATES.

A MARKOVIAN AUTOMATON $\mathcal{A}$ is a right action of the free monoid X^* (the words made from X) on Q, with a INITIAL PROBABILITY VECTOR $\gamma_0 \in [0,1]^Q$ (i.e., $\sum_{q \in Q} \gamma_0(q) = 1$) and a matrix $M \in [0,1]^{Q \times X}$ such that every ROW is a probability vector (error in the notes!) i.e., $\forall\, q \in Q$ $\sum_{a \in X} M(q,a) = 1$.

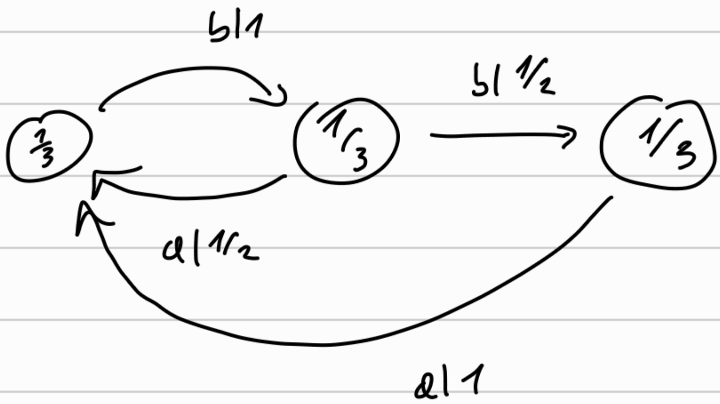
SCHEMATICALLY: It is a directed labelled graph with:

- Q as vertex set, with labels given by γ_0.
- there is an edge between q_1 and q_2 iff $\exists\, a \in X$ st $q_1 a = q_2$ with $M(q_1, a) > 0$, and the label is $M(q_1, a)$.

EXAMPLE:

Q = three states. $X = \{a, b\}$ $\gamma_0 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$

$$M = \begin{array}{cc} a & b \\ \left(\begin{array}{cc} 0 & 1 \\ 1/2 & 1/2 \\ 1 & 0 \end{array}\right) \end{array}$$



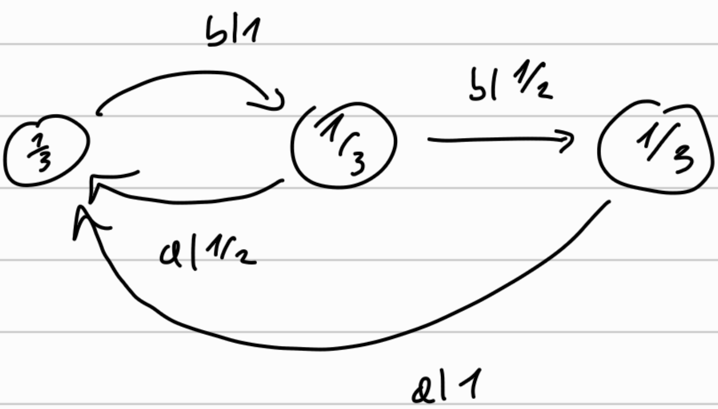
Def: Let $\mathcal{A} = (X, Q, \gamma_0, M)$ be a Markovian automaton.

For each n, the probability induced by $\mathcal{A}$ on R_n is defined by:

$$\mathbb{R}_n(x_1 x_2 \cdots x_n) = \sum_{q \in Q} \gamma_0(q)\, M(q, x_1)\, M(q \cdot x_1, x_2) \cdots M(q \cdot x_1 \cdots x_{n-1}, x_n)$$

For every state $q \in Q$, I act on it with the word, keeping track of the states I end up with and the probabilities of getting there

EXAMPLE: $X = \{a, b\}$, $Q = \{q_1, q_2, q_3\}$



abab

state q_1	a does not go anywhere		$M(1, a) = 0$		0
					$+$
state q_2	$a: q_2 \to q_1$	with prob $1/2$			
	$b: q_1 \to q_2$	" " 1		$1/3 \cdot 1/4 =$	$1/12$
	$a: q_2 \to q_1$	$1/2$			
	$b: q_1 \to q_2$	1			
					$+$
state q_3				$1/3 \cdot 1/2 =$	$1/6$
					$=$
					$1/4$

ANOTHER EXAMPLE: We are interested in reduced words in the free group:

$A = \{s_1, \dots, s_r\}$ r generators

$\tilde{A} = A$ and inverses, $\tilde{A} = \mathcal{Q}$

the transitions are given by

$a : b \longrightarrow a$

We put a non zero entry in M iff $b \neq a^{-1}$ (we want reduced words)

If every non zero entry of M is $\frac{1}{2r-1}$ and $\gamma_0 = \left(\frac{1}{2r}, \dots \frac{1}{2r}\right)$ we get the uniform distribution.

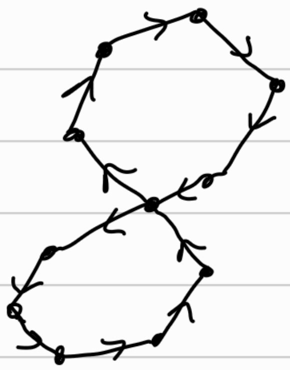
Rmk: By changing π_n, this includes few relators and density models.

With these automata we can get prefix heavy distributions.

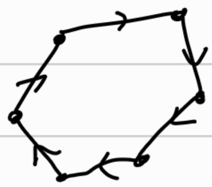
Theorem (Bassino et al) [Thm 4.3]
Let $\mathcal{A}$ be a Markovian automaton.
If $\mathcal{A}$ does not have a cycle with probability 1, then the $\mathbb{R}_n$ it induces are prefix heavy with parameters (C, α) explicitly computable.

Sketch of proof

A CYCLE is a path in $\mathcal{A}$ that begins and ends on the same state.
A CYCLE is ELEMENTARY if it does not visit a state more than once.



a cycle, not elementary



a simple cycle.

Let δ the maximum probability of an elementary cycle in $\mathcal{A}$, and let ℓ be the maximum length of such a cycle.

Let $\gamma(q, u)$ = probability of a cycle u at state q.

We have $\gamma(q,u) \leq \delta^{\frac{|u|}{\ell}}$

For any word w (thus any path) we can write it as a sequence of cycles and at most $|Q|$ segments. thus

$$\gamma(q, w) \leq \delta^{\frac{|w| - |Q|}{\ell}}$$

$$\mathbb{R}_u(w) \leq |Q| \, \delta^{\frac{|w| - |Q|}{\ell}}$$

Putting $\quad C = |Q|\, \delta^{-\frac{|Q|}{\ell}} \qquad \alpha = \delta^{1/\ell}$

$$\mathbb{R}_n(P(uv)) = \mathbb{R}_{|uv|}(uv) = \sum_{q \in Q} \gamma_0(q)\gamma(q,u)\,\gamma(q \cdot u, v)$$

$$\leq \Big(\sum_{q \in Q} \gamma_0(q)\gamma(q,u) \Big)\, C \alpha^{|v|}$$

$$= \mathbb{R}_{|u|}(u)\, C\alpha^{|v|} = \mathbb{R}_n(P(u))\, C\alpha^{|v|}$$

$\square$

α DENSITY MODEL

A Markovian automaton defines a variant of the classical density model.

Def (α - density model):
Let A be a Markovian automaton, $0 < \alpha < 1$.
Let $\mathbb{R}_n$ be the sequence of probabilities induced by A and Π_n be the measure with support on the
$\lfloor \alpha^{dn} \rfloor$ - tuple $(u, \dots, u)$.

The resulting probabilities $\mathbb{P}_n$ give the α - density model.

Rmk:
If $\alpha = \frac{1}{2r-1}$ and A is as the last example, we get the usual density model.

Theorem (4.3)
If A is a Markovian automaton without cycles of probability 1 ((C, α) - prefix heavy), then:

1) In the α - density model at density $d < 1/4$ a tuple of reduced words exp. generically has the central tree property.

2) At α - density $d < 1/6$, a tuple of reduced words exp. generically generates a malnormal subgroup.

Sketch of proof for 1) (Apparently wrong...)

Let P be the probability that a α^{dn}-tuple $\vec{h}$ does NOT satisfy the CTP.

No CTP $\Rightarrow$ $\exists$ a word of length $t = \frac{1}{2}n$ that occurs as prefix of h_i or h_i^{-1} and h_j or h_j^{-1} for some $i < j$

then $$P \leq 4 \sum_{i<j} \sum_{w \in R_t} \mathbb{R}_n(P(w))^2$$

Since $\mathbb{R}_n$ is prefix heavy, $\mathbb{R}_n(P(w)) \leq C\alpha^t$ for $w \in R_t$

$$\sum_{w \in R_t} \mathbb{R}_n(P(w))^2 \leq \underbrace{\sum_{R_t} \mathbb{R}_n(P(w))}_{=1} \cdot C\alpha^t$$

So $$P \leq 4 \sum_{i<j} C\alpha^t = 4C\,\alpha^{(2d-1/2)n}$$

So this decreases exp if $2d - 1/2 > 0$

$\underline{d > 1/4}$???

$\boxed{?}$