Stationary measures for Variable Length Markov Chains

Joint work:

P. Cénac (université de Bourgogne, CNRS, IMB, Dijon)
B. Chauvin (université Paris-Saclay, CNRS, LMV, Versailles)
F. Paccaut (université de Picardie Jules Vernes, CNRS, LAMFA, Amiens)
N. Pouyanne (université Paris-Saclay, CNRS, LMV, Versailles)

AofA, CIRM (Luminy), 28 june 2019

“Variable Length Markov Chain”

Blackboard

Take an alphabet $\mathscr{A}$ (finite set, $\mathscr{A} = \{0, 1\}$, say)

• A random chain (or a probabilistic source) on $\mathscr{A}$: a sequence $X_0, X_1, X_2 \cdots$ of $\mathscr{A}$-valued r.v.

You may suppose that the chain has been alive since the dawn of time: one gets $\cdots X_{-2}, X_{-1}, X_0, X_1, X_2 \cdots$

As you are not used to, write this sequence by concatenation $U_n = X_n X_{n-1} X_{n-2} \cdots$: U_n is a right-infinite word

So, start with a random $U_0 = X_0 X_{-1} X_{-2} \cdots$

Add a random letter on the left: $U_1 = X_1 X_0 X_{-1} \cdots$

And another one $U_2 = X_2 X_1 X_0 \cdots$ and so on, so that $\boxed{U_{n+1} = X_{n+1} U_n}$

• The chain is Markov whenever $\mathbf{P}\left(U_{n+1} = \alpha U_n | U_n\right)$ is a function of X_n (one letter from the past is enough to predict the future)

It is Markov of order 2 whenever $\mathbf{P}\left(U_{n+1} = \alpha U_n | U_n\right)$ is a function of $X_n X_{n-1}$ (two letters from the past is enough to predict the future)

Idem for a Markov chain of order $m \geqslant 1$

• Markov chains of variable order (VLMC): $\mathbf{P}\left(U_{n+1} = \alpha U_n | U_n\right)$ is a function of a (not bounded) prefix of U_n.

Details come later on.

• Examples: last run of 0, last run, last $0^p 1^q$, avoid pattern (21) with 3 letters

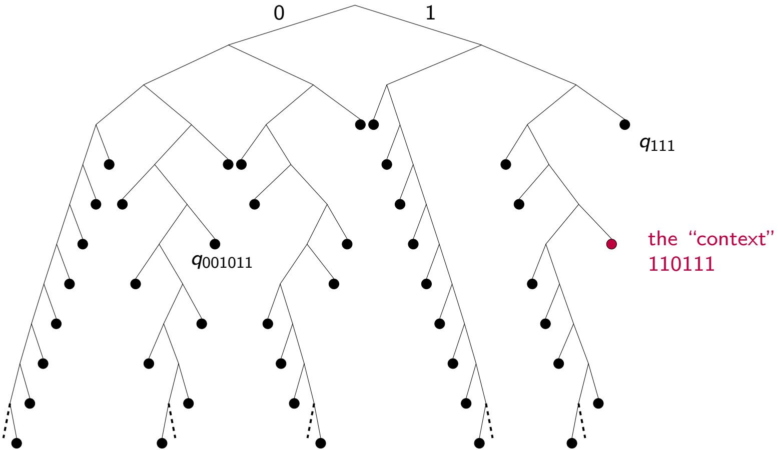
How one defines a VLMC?

Take:

(i) A saturated tree on the alphabet (*context tree*)

(ii) On any leaf c, a probability measure q_c on the alphabet $\{0, 1\}$

[(iii) For the initial U_0, a distribution on the set of right-infinite words]

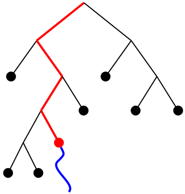


How one defines a VLMC: hanging the right-infinite words

Blackboard

Take a probabilised context tree (picture). Conditionally to the right-infinite word U_n, let's see how one defines the transition probability $\mathbf{P}(U_{n+1} = \alpha U_n | U_n)$:

- $\star$ take the right-infinite word U_n
- $\star$ hang it by its first letter
- $\star$ insert it at the root of the tree
- $\star$ look through which leaf it goes out.



$U_n = 01011100\cdots$

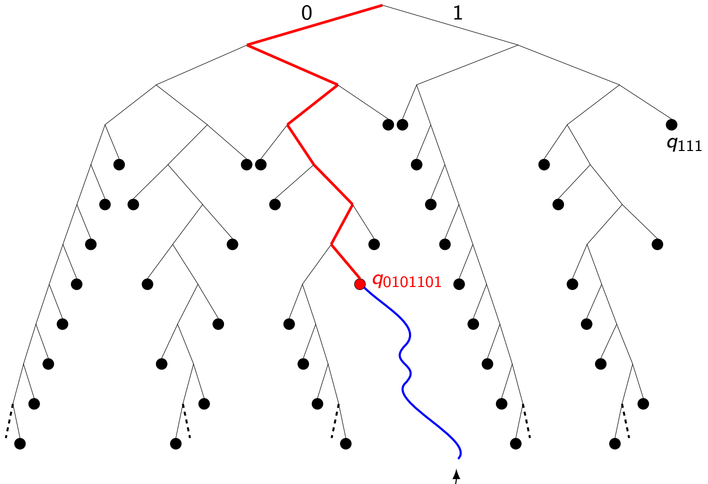
$\mathrm{pref}(U_n) = 0101$

Denote this leaf by $\mathrm{pref}(U_n)$.

Then, for every $\alpha \in \{0,1\}$, take

$$\mathbf{P}(U_{n+1} = \alpha U_n | U_n) = q_{\mathrm{pref}(U_n)}(\alpha)$$

This defines a Markov process on right-infinite words.

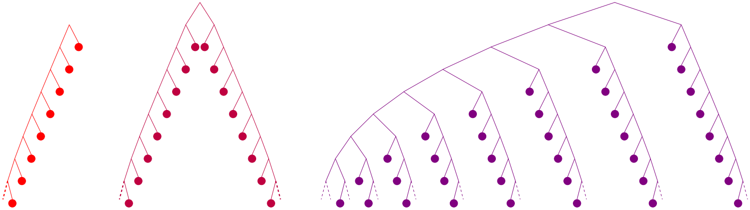


Insert $U_n = 0101101110\cdots$

$\mathrm{pref}\,(U_n) = 0101101$

Examples

Last run of 0: left comb
Last run (of 0 or 1): double comb
Last pattern of the form 0^p1^q: left comb of right combs



The question

Does such a Markov chain admit an invariant probability measure?

In other words, is it possible to choose the law of U_0 so that every U_n has the same distribution?

Remark: the state space is uncountable.

Assumptions

(i) The context tree has an at most countable set of infinite branches

[For instance, the complete infinite tree is forbidden]

(ii) $q > 0$ (*non-nullness*)

[Remark: when the context tree is finite, say of height h, the VLMC is a standard (economical) Markov chain of order h (PF).]

Combinatorial definition: the LIS of a finite word

Take a context tree.
Take a finite non-empty word w and write it (uniquely) as

$$w = p\alpha s$$

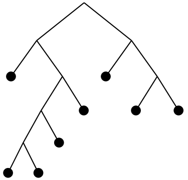
where
(i) p is a prefix of w, possibly empty
(ii) α is a letter
(iii) s is *the Longest Internal (proper) Suffix* of w.

Shortly said, s is the *LIS* of w, possibly empty

[The suffix αs is also named the *alpha-LIS* of the world w.]

Examples of LIS and α-LIS computations

Blackboard

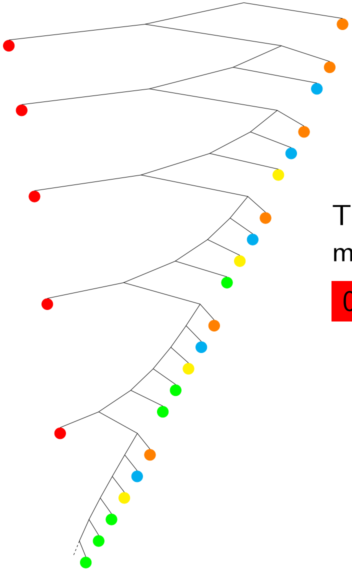


Take $w = 01101001$.
Compute its alpha-LIS: 001

Alpha-LIS of interest for the sequel: the context alpha-LIS

context	*alpha − LIS*
00	00
01000	00
01001	001
0101	101
011	011
10	10
110	10
111	111

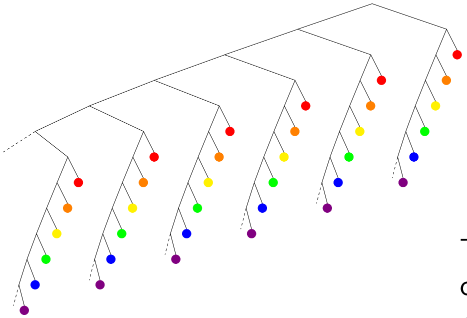
Tree examples with their context alpha-LIS (1)



The 1-arithmetical tree has finitely many context alpha-LIS:

00, 001, 001001, 101 and 1

Tree examples with their context alpha-LIS (2)



The left comb of left combs has an infinite set of context alpha-LIS: the 10^n1, $n \geqslant 0$.

Cascades

Blackboard

Take $w = p\alpha s = \beta_1\beta_2\beta_3\cdots\beta_\ell\alpha s$ where s is the LIS of w.

The *cascade* of w is defined as

$$\begin{aligned}\mathrm{casc}(w) = \quad & q_{\mathrm{pref}(\beta_2\cdots\beta_\ell\alpha s)}(\beta_1) \\ & \times q_{\mathrm{pref}(\beta_3\cdots\beta_\ell\alpha s)}(\beta_2) \\ & \vdots \\ & \times q_{\mathrm{pref}(\alpha s)}(\beta_\ell)\end{aligned}$$

If π is a stationary probability measure on the set $\mathscr{R}$ of right-infinite words, then

$$\pi\left(w\mathscr{R}\right) = \mathrm{casc}(w) \times \pi\left(\alpha s\mathscr{R}\right)$$

This is the *cascade formula*.

The Q matrix

Blackboard

Define the square matrix Q, indexed by the set of context alpha-LIS:

$$Q_{\alpha s, \beta t} = \sum_{\substack{c:\ context \\ c = t \cdots = \cdots [\alpha s]}} \operatorname{casc}(\beta c)$$

Remarks:
(i) size of Q: the index set is at most countable (the set of context is)
(ii) Q is defined from the VLCM data (context tree and q's)

The cascade series'

Blackboard

If αs is a context alpha-LIS, the cascade series of αs is

$$\boxed{\kappa_{\alpha s} = \sum_{c = \cdots [\alpha s]} \mathrm{casc}(c) \in [0, +\infty]}$$

When all the $\kappa_{\alpha s}$ are finite, one says that *the cascade series converge*.

Remark: the convergence of the cascade series is sufficient to ensure that Q's entries are finite.

Main result

Notation: $\mathscr{S}$ = the set of context α-lis

Remember: $\kappa_{\alpha s} = \sum_{c=\cdots[\alpha s]} \mathrm{casc}(c)$ is the cascade series of $\alpha s \in \mathscr{S}$

Theorem (CCPP, 2018) *Take a non-null VLMC, named U.*

(i) If U gets some invariant probability measure, then all the cascade series converge.

(ii) Assume that the cascade series converge. Then

$$\pi \mapsto \left(\pi\left(\alpha s \mathscr{R}\right)\right)_{\alpha s \in \mathscr{S}}$$

is a bijection between:

- the U-invariant probability measures π

- the left-fixed vectors $\left(v_{\alpha s}\right)_{\alpha s \in \mathscr{S}}$ of Q that satisfy

$$\sum_{\alpha s \in \mathscr{S}} v_{\alpha s} \kappa_{\alpha s} = 1.$$

Pleasant case: whenever $\mathscr{S}$ is finite (Q is finite-dimensional).

Stable trees

A context tree is said *stable* whenever it is stable under the shift on words (or after removal of the inital, or after decapitation of words).

Theorem (CCPP, 2018)
(i) When the context tree is stable, the matrix Q is row-stochastic and irreducible

(ii) Assume that a VLMC is defined over a stable context tree, and that the set of alpha-LIS is finite. Then,

$$\exists! \text{ invariant probability} \iff \text{the cascade series converge}$$

Last slide of AofA'19, enfin

① Example:
Alphabet $=\{0,1,2\}$
(Conditional) law of U_{n+1} depends on the longest increasing sequence of U_n
$\rightsquigarrow$ Stable tree, 3 context alpha-LIS: no mystery!

② Non stable trees: lots of open questions.
A suivre...

Merci !