A “finite state” version of the Kraft-McMillan theorem

F. Bassino, M.-P. Béal, D. Perrin

Institut Gaspard-Monge
Université de Marne la Vallée. France

http://www-igm.univ-mlv.fr/~{bassino,beal,perrin}

Outline

- The Kraft-McMillan theorem
- The question of enumerative sequences of leaves
- Super-state automata
- The question of enumerative sequences of nodes

The Kraft McMillan theorem

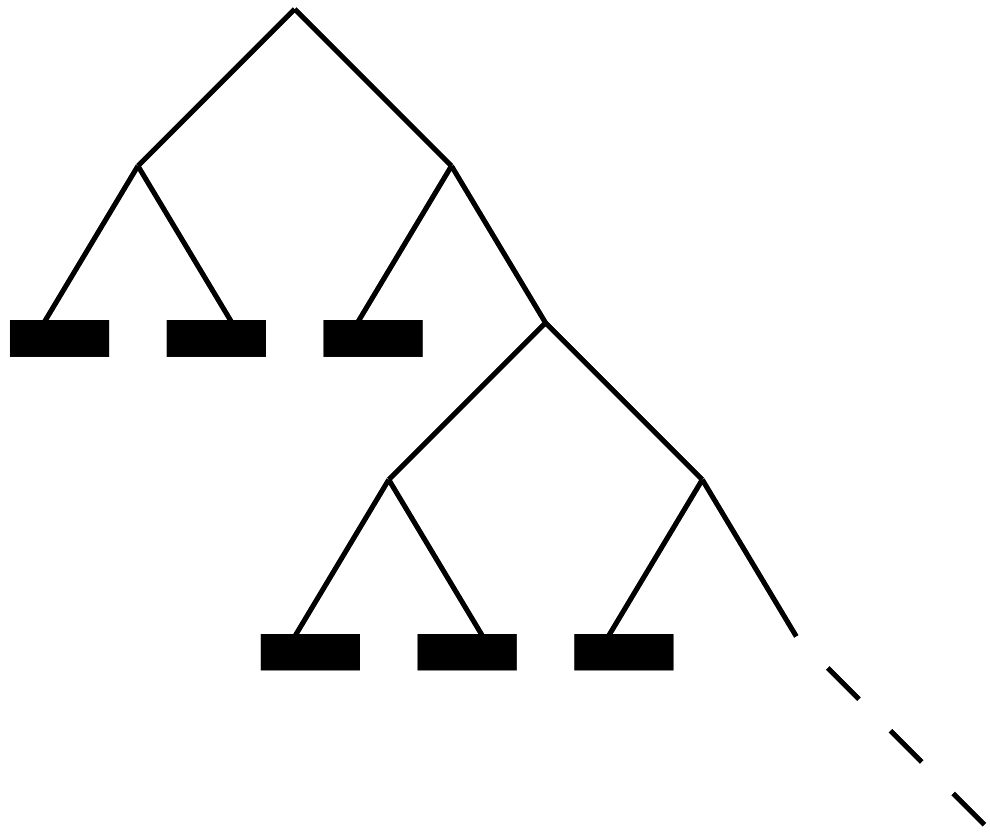
A sequence of nonnegative integers $(s_n)_{n\geq 0}$ is the enumerative sequence of leaves of a k-ary tree (resp. a complete k-ary tree) iff

$$\sum_{n\geq 0} s_n k^{-n} \leq 1 \text{ (resp. } = 1)$$

$$s(z) = \sum_{n\geq 0} s_n z^n, \; s(\frac{1}{k}) \leq 1$$

s_n is the number of leaves at height n.

- compression coding schemes
- coding for constrained channels
- construction of prefix codes



Tree associated to $3z^2(z^2)^* = 3z^2/(1-z^2)$.

Problem

Characterize the enumerative sequences of leaves of k-ary rational trees.

rational tree: a tree which has a finite number of non isomorphic subtrees.

k-ary : each node of the tree has at most k sons.

Necessary conditions

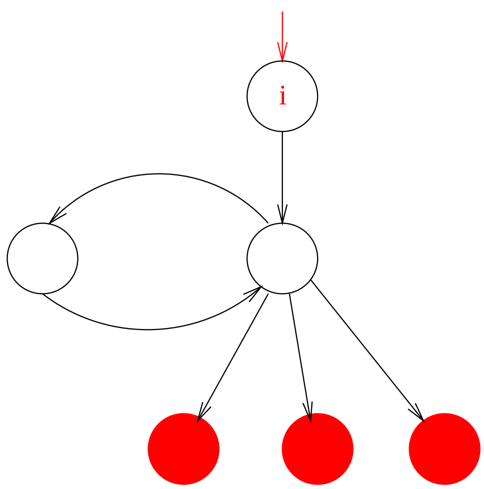
- Kraft inequality: $\sum_{n \geq 0} s_n k^{-n} \leq 1$
- $\mathbb{N}$-rational sequence: s_n is the number of paths of length n going from an initial state to a final state in an automaton (a finite state machine).

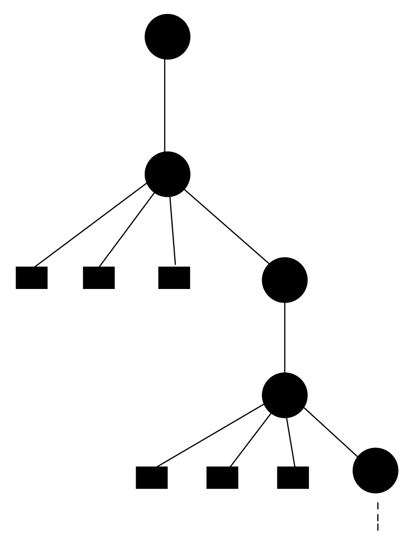
In general, the tree obtained with the Kraft-McMillan construction is not necessarily rational.

Another construction

We begin with a normalized automaton that represents the sequence. We "develop" the automaton and get a rational tree. This tree is not necessarily k-ary.

$$s(z) = 3z^2(z^2)^* = 3z^2/(1 - z^2)$$





Result

Any $\mathbb{N}$-rational sequence satisfying the Kraft inequality is the enumerative sequence of leaves by height of a k-ary rational tree.

Perron-Frobenius

Let M be a nonnegative irreducible matrix. Its spectral radius λ is a simple eigenvalue of M. There is a positive eigenvector $\mathbf{v}$ such that $M\mathbf{v} = \lambda\mathbf{v}$.

If $\lambda < k$, where k is a positive integer, there is an approximate eigenvector, that is a positive vector $\mathbf{v}$ with $M\mathbf{v} \leq k\mathbf{v}$.

Computation of an a. eigenvector

Franaszek's algorithm

begin

$\mathbf{v} := (1, 1, \ldots, 1)^t$;

repeat

begin

$\mathbf{w} := \mathbf{v}$;

$\mathbf{v} := \max\left\{\mathbf{w}, \lceil \frac{1}{k} M\mathbf{w} \rceil\right\}$;

end

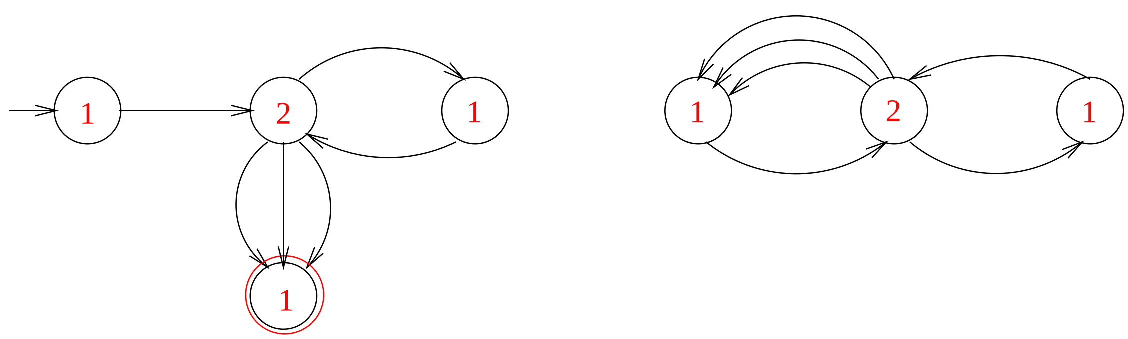
until $\mathbf{v} = \mathbf{w}$;

end

From the series to the (a.) eigenvector

Let s be the sequence recognized by a normalized automaton (G, i, F). Let $\bar{G}$ be the strongly connected graph obtained from G by identification of initial and final states.

$$S(z) = 3z^2(z^2)^* = \frac{3z^2}{1-z^2}$$



$$M = \begin{pmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad M\mathbf{v} = 2\mathbf{v}$$

λ

- does not depend on the representation G. ($\frac{1}{\lambda}$ is the minimal modulus of the poles of the series $\frac{1}{1-s(z)}$).

$$\sum_{n \geq 0} s_n k^{-n} < 1 \text{ ssi } \lambda < k$$

$$\sum_{n \geq 0} s_n k^{-n} = 1 \text{ ssi } \lambda = k$$

We take a k-approximate eigenvector $\mathbf{v}$: $M\mathbf{v} \leq k\mathbf{v}$.

Super-state automata

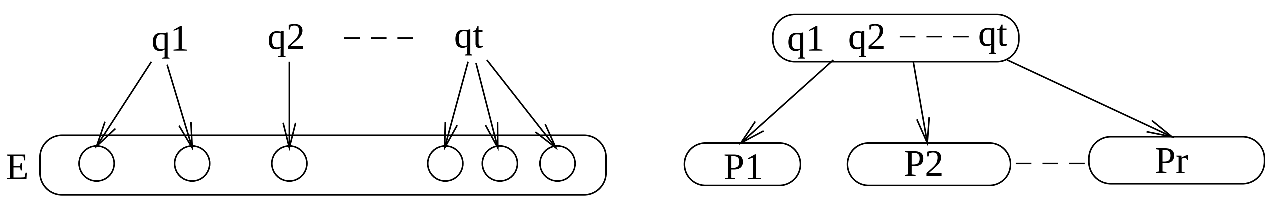
$\mathcal{A} = (Q, F) \longrightarrow \mathcal{B} = (Q', F')$

$+$ valuation v of the states $+$ integer m

$Q' = \{(q_1, q_2, \ldots, q_t) \mid q_i \in Q, 1 \leq t \leq m\}$

set of unordered t-uples.

$v((q_1, q_2, \ldots, q_t)) = \sum_{j=1}^{t} v(q_j)$



$P_1, P_2, \ldots, P_r$ is a partition of E.

The partition is such that all parts but one have a valuation dividable by m.

Weight lemma or pigeon-hole

[Brian Marcus [2]]
Let $v_1, v_2, \ldots, v_m$ *be integers. There is a subset* S *of* $\{1, 2, \ldots, m\}$ *such that* $\sum_{q \in S} v_q$ *is dividable by* m*.*

Paul Erdős attributes this nice application of the pigeon-hole principle to Marta Sved and Andrew Vázsonyi in [1] p. 125.

References

[1] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*, Springer, 1998.

[2] Brian Marcus. Factors and extensions of full shifts *Monats. Math.*, 88:239-247, 1979

proof:

$$m \begin{cases} v_1 \\ v_1 + v_2 \\ v_1 + v_2 + v_3 \\ \vdots \\ v_1 + v_2 + \ldots + v_m \end{cases}$$

either there is a p such that, modulo m,

$$v_1 + v_2 + \ldots + v_p \equiv 0$$

or there are p, r such that, modulo m,

$$v_1 + v_2 + \ldots + v_p \equiv v_1 + v_2 + \ldots + v_r$$

Proof

$s \longrightarrow G \longrightarrow \bar{G} \longrightarrow \mathbf{v} : M\mathbf{v} \leq k\mathbf{v}\ (\lambda \leq k)$

One takes $\mathbf{v}$ as valuation and $m = v(i)$.

Let u be a super-state. One defines:

$$w(u) = \lceil v(u)/m \rceil$$

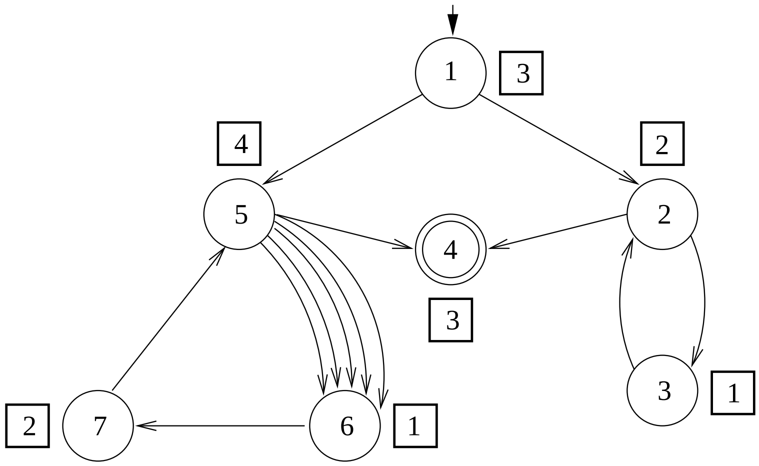
If u has t outgoing edges that arrive in $u_1, \ldots u_t$,

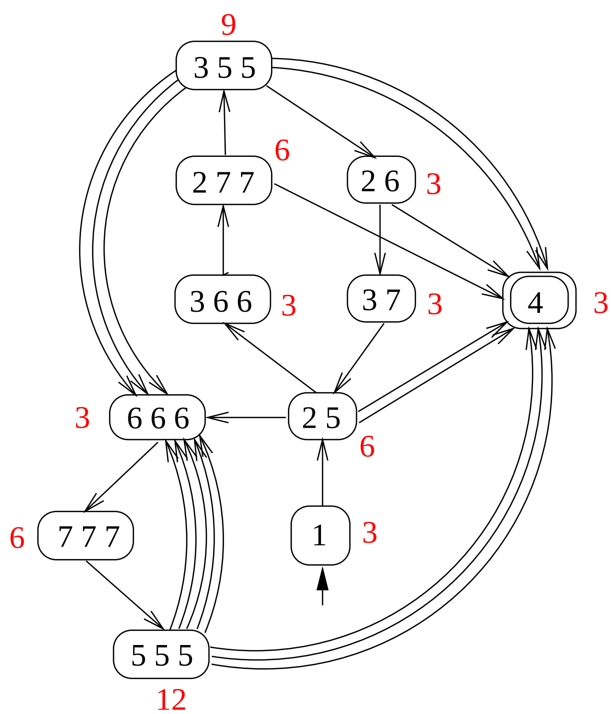
$$\sum_{j=1}^{t} v(u_j) \leq kv(u)$$

$$\sum_{j=1}^{t-1} v(u_j)/m + \lceil v(u_t)/m \rceil \leq k\lceil v(u)/m \rceil$$

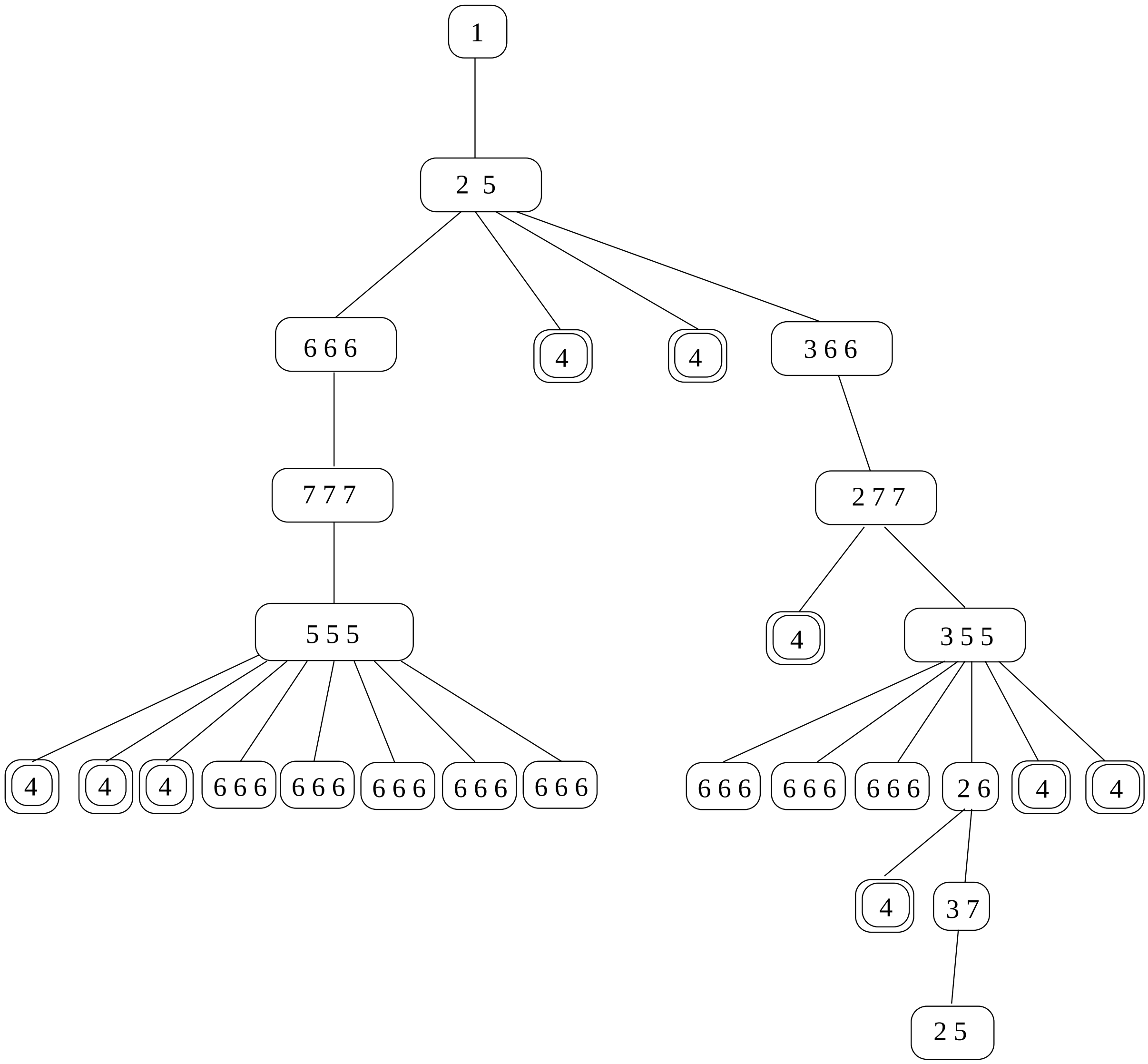
$$\sum_{j=1}^{t} w(u_j) \leq kw(u) \text{ et } w(i) = 1$$

$$s(z) = \frac{z^2}{1-z^2} + \frac{z^2}{1-5z^2}$$

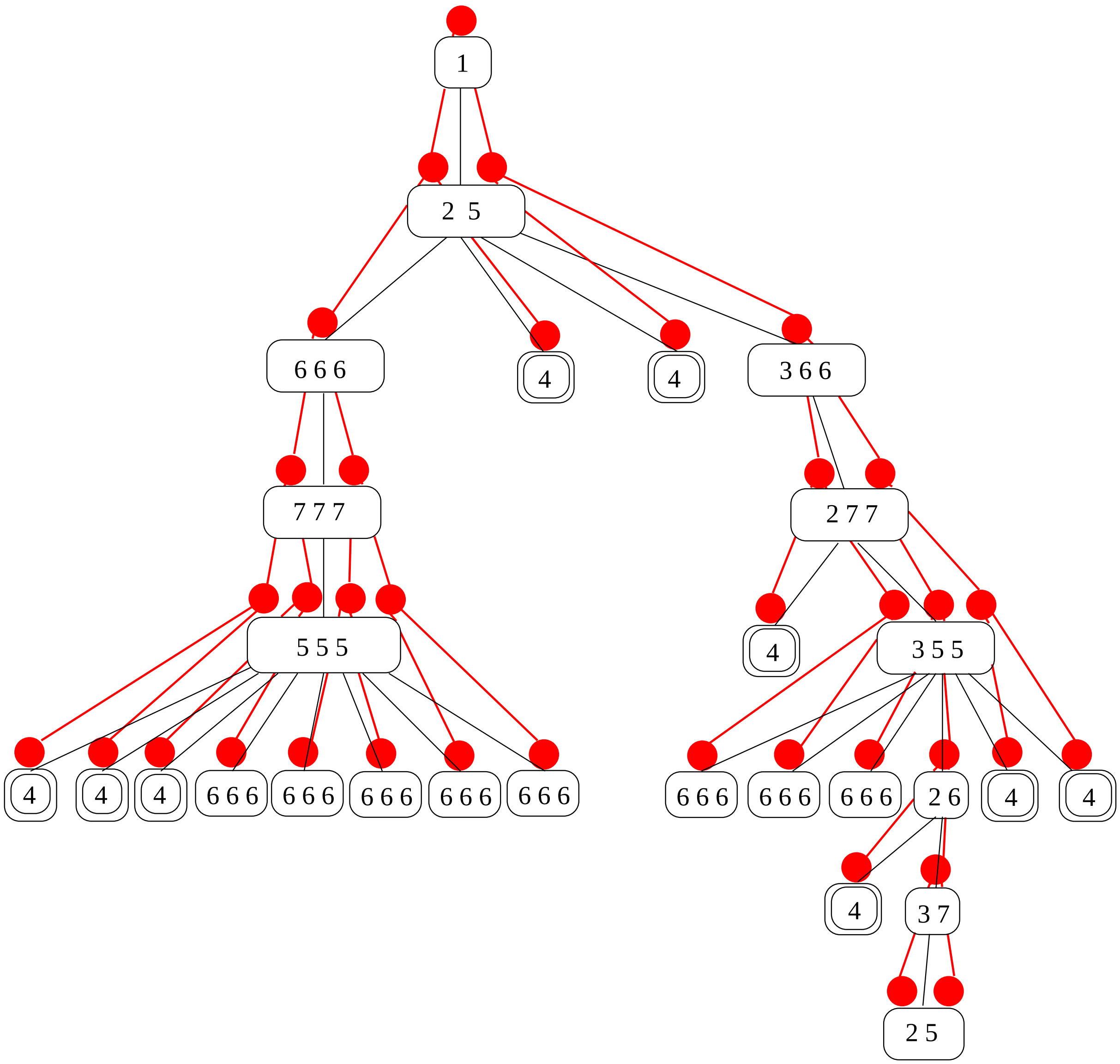




Super-state automaton with the valuations.



The developed automaton.



The number of small red balls above a super-state u is $w(u) = \lceil \frac{v(u)}{m} \rceil$ (here $m = 3$).

The sequences of nodes

An $\mathbb{N}$-rational sequence $(t_n)_{n\geq 0}$ is the enumerative sequence of nodes by height of a k-ary rational tree iff:

- $t_0 = 1$
- $\forall n \geq 1, t_{n+1} \leq kt_n$
- the convergence radius of the series t is $> \frac{1}{k}$
- the series $s(z) = t(z)(kz - 1) + 1$ is $\mathbb{N}$-rational.

There are $\mathbb{N}$-rational series that satisfy the first three conditions but not the last one.

References

[1] F. Bassino, M.-P. Béal, and D. Perrin. Super-state automate and rational trees. In *LATIN'98*, LNCS 1380, pages 42–52. Springer, 1998.

[2] F. Bassino, M.-P. Béal, and D. Perrin. Enumerative sequences of leaves and nodes in rational trees. *Theoret. Comput. Sci.*, 1997. (to appear).

[3] F. Bassino, M.-P. Béal, and D. Perrin. Enumerative sequences of leaves in rational trees. In *ICALP'97*, LNCS 1256, pages 76–86. Springer, 1997.

[4] D. Perrin. Arbres et séries rationnelles. *C.R.A.S. Paris, Série I*, 309:713–716, 1989.

[5] D. Perrin. A conjecture on rational sequences. In R. Capocelli, editor, *Sequences*, pages 267–274. Springer-Verlag, 1990.