Languages of Markov Chains

Based on:

S. Akshay, Blaise Genest, Bruno Karelovic, Nikhil Vyas.

On Regularity of unary Probabilistic Automata STACS 2016

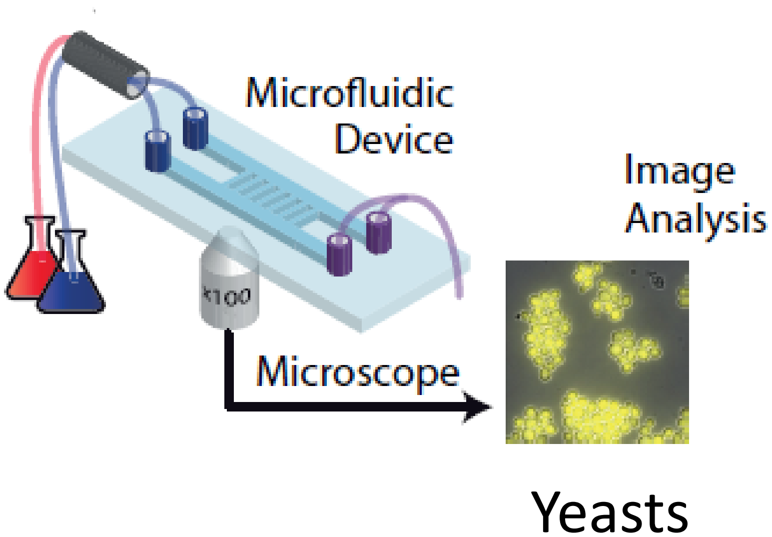
Manindra Agrawal, S. Akshay, Blaise Genest, P.S. Thiagarajan.

Approximate Verification of the Symbolic Dynamics of Markov Chains JACM'15-LICS'12

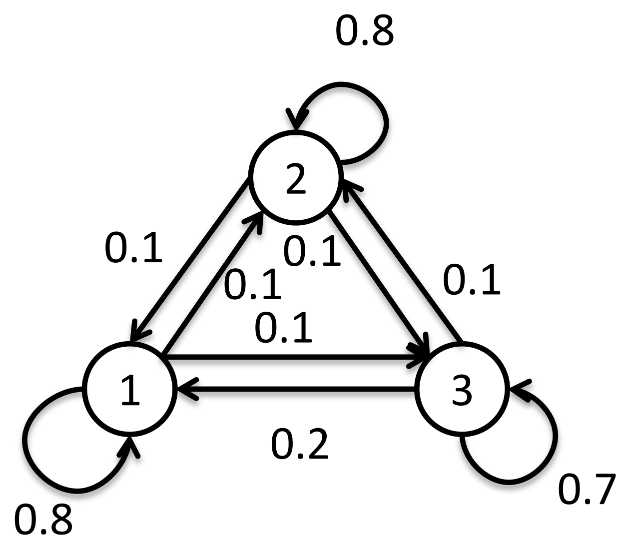
Overview of the Talk

- Biological Motivation
- Trajectories of Markov Chains
- Languages of Markov Chains
- Approximations

Population of cells



Stochastic
Model



Setting: Many yeasts. Simplistic model
Each yeast can be in one of 3 states (1 high, 2 med,3 low concentr. of X)

Experiments: Percentage of yeasts going from state S to state S’
=> Chance for a yeast to go from state S to state S’ (after 5 min).

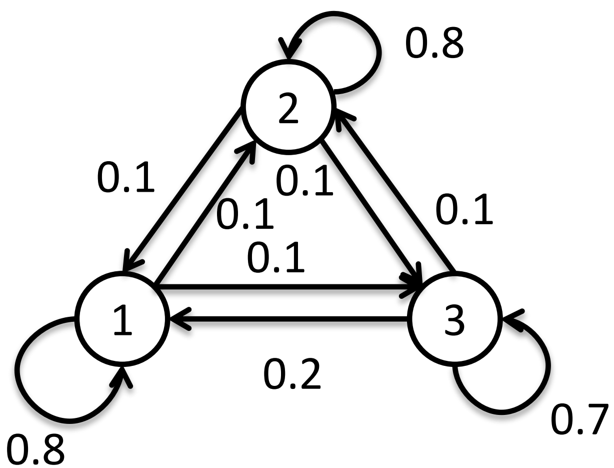
Image analysis: proportion in state 1 (high concentration of X => marker).

Continuous Abstraction

Non Continuous: Nathalie'sTalk tomorow.

Assume enough yeasts => Proportion in state 1,2,3.

Initial Proportion of cells : $P_{init} = \begin{pmatrix} - \\ \\ - \end{pmatrix}$ through image analysis



$$M_{yeast} = \begin{pmatrix} 0.8 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.7 \end{pmatrix}$$

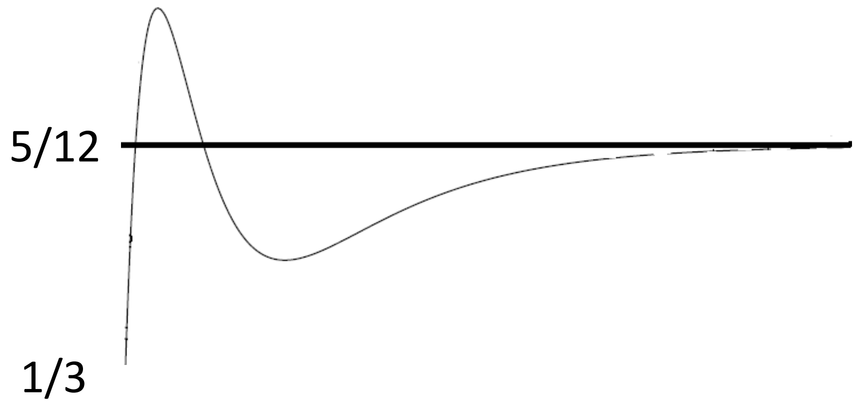
$P_{5min} = M_{yeast}\, P_{init}$ $P_{10min} = M^2_{yeast}\, P_{init}$ ....

Deterministic concrete trajectory from a given P_{init}

Symbolic Trajectory

Expermiments (image analysis):

first less than 5/12 of yeasts in state 1.
some time later more than 5/12 of yeasts in state 1.
then eventually, less than 5/12 of yeasts in state 1.



We set up Threshold=5/12

Below threshold: B
Above (or equal) threshold: A => finite alphabet {A,B}

So we observed $B^{n1}A^{n2}B^{\omega}$ Symb. Trajectory = infin. word on {A,B}

Language of Markov Chain: set of trajectories from Init={P_{init} | $x \in [0,2/3]$}

Quantitative Question

Is $B^{n1}A^{n2}B^{\omega}$ in the language of the Markov Chain for some n1,n2 ?

If yes, for which initial proportion, for which n1,n2?

i.e. for which subset of s in $\quad Init = \{ \begin{pmatrix} - \\ \\ - \end{pmatrix} \mid x \quad [\; -] \quad ?$

Looks like a verification Question.
Use algorithm for solving PCTL* questions on Markov Chains ?

Cannot be modeled with PCTL* [Beauquier Rabinovitch Slissenko CSL'02]

Skolem Problems

Actually, even with a unique initial configuration $P_{init} = \begin{pmatrix} & - & \\ & & \\ - & & \end{pmatrix}$

Trajectory of a Markov chain from P_{init} is B^{ω} ?

as hard as Skolem (question on linear rec. seq. (e.g. Fibonacci))
[Akshay, Antonopoulose,Ouaknine, Worrel, IPL'15]

Decidability? Open for > 40 years.
Decidable for <6 states
If dec. for 18 states, major breakthrough in diophantines approximations

Simple Markov Chains

Simple: Every Eigen value of Markov Chain has multiplicity 1.

-> Markov Chain is diagonalizable.

For simple markov chains,

Trajectory of a Markov chain from some P_{init} is B^{ω} ?

=> decidable for 10 states.

=> for more than 25 states, breakthrough in Diophantines approx

Decidable if Trajectory of a simple Markov chain from some P_{init} is wB^{ω} for some finite word w (ultimate positivity)

[Ouaknine Worrell ICALP'14 (best paper) & ICALP'14]

Eigen Basis

Simple:Markov Chains are diagonalizable.

$$M_{yeast} = \begin{pmatrix} 0.8 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.7 \end{pmatrix}$$

Express P_{init} in the eigen vector basis

$$M \cdot \begin{pmatrix} 5/12 \\ 1/3 \\ 1/4 \end{pmatrix} = \underline{1} \begin{pmatrix} 5/12 \\ 1/3 \\ 1/4 \end{pmatrix} \; ; \; M \cdot \begin{pmatrix} 5/12 \\ -5/12 \\ 0 \end{pmatrix} = \underline{0.7} \begin{pmatrix} 5/12 \\ -5/12 \\ 0 \end{pmatrix} \; ; \; M \cdot \begin{pmatrix} 5/12 \\ 0 \\ -5/12 \end{pmatrix} = \underline{0.6} \begin{pmatrix} 5/12 \\ 0 \\ -5/12 \end{pmatrix}$$

$$P_{init} = \begin{pmatrix} 1/3 \\ 1/4 \\ 5/12 \end{pmatrix} = \begin{pmatrix} 5/12 \\ 1/3 \\ 1/4 \end{pmatrix} + \frac{1}{5} \begin{pmatrix} 5/12 \\ -5/12 \\ 0 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 5/12 \\ 0 \\ -5/12 \end{pmatrix}$$

$$M^n P_{init} = \begin{pmatrix} 5/12 \\ 1/3 \\ 1/4 \end{pmatrix} + \frac{1}{5}\, 0.7^n \begin{pmatrix} 5/12 \\ -5/12 \\ 0 \end{pmatrix} - \frac{2}{5}\, 0.6^n \begin{pmatrix} 5/12 \\ 0 \\ -5/12 \end{pmatrix}$$

$M^n P_{init}\ [1] \geq$ — iff $-0.7^n \geq -0.6^n$

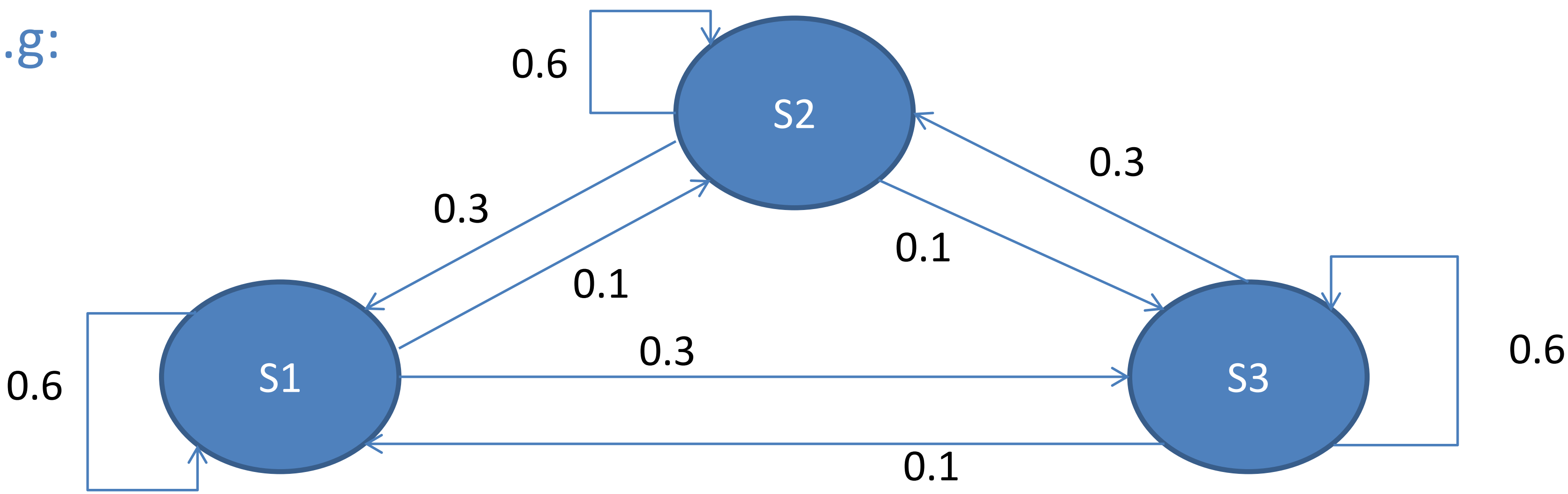
Symbolic Trajectory from P_{init}: $B^k A^\omega$

General Simple Markov Chains

$M^n P_{init}$ [1] $\geq \tau$ iff $a_0 + a_1 \sigma_1^n + \ldots + a_k \sigma_k^n \geq 0$

Trajectories are not always ultimately periodic, even for simple M

e.g:



$M^n P_{init}$ [0] $\geq 1/3$? with $P_{init}(S1) = P_{init}(S2) = 1/4$.

Reason: eigen values: 1, —— ——

Roots of real numbers

If eigen values are roots of real numbers,
then trajectories are ultimately periodic

$M^n P_{init}$ [1] $\geq \tau$ iff $a_0 + a_1 \sigma_1^n + \ldots + a_k \sigma_k^n \geq 0$

Let l_k with $\sigma_k^{l_k}$ is positive real,
Let $L = \mathrm{lcm}(l_k)$ and $\rho_k = \sigma_k^L$

$M^{Ln} P_{init}$ [1] $\geq \tau$ iff $a_0 + a_1 \rho_1^n + \ldots + a_k \rho_k^n \geq 0$

→ Eventually constant (dominant factor) e.g.: w A^ω

→ Ultimately periodic of period L
ex: L=5, trajectory w' $(A\ B\ B\ A\ B)^\omega$

Sum up

Property of eigenvalues of Markov chain	Ultimately periodic traj.
Distinct, positive real numbers	✓ even ult. constant
Distinct, roots of real numbers	✓ Decidability!
Distinct	× Decidability ?

but positivity/equality approximable

[Chadha Kini Viswanathan QEST’14]

What about languages?

set of trajectories from e.g. $\text{Init} = \{ \begin{pmatrix} - \\ \\ - \end{pmatrix} \mid x \quad [\quad -]$

Init needs to be a polytope.

Language: not that simple

M_{yeast}: eigen values are positive real numbers: 1, 0.7, 0.6

$$\begin{pmatrix}1/3\\1/4\\5/12\end{pmatrix}=\begin{pmatrix}5/12\\1/3\\1/4\end{pmatrix}+\frac{1}{5}\begin{pmatrix}5/12\\-5/12\\0\end{pmatrix}-\frac{2}{5}\begin{pmatrix}5/12\\0\\-5/12\end{pmatrix}$$

Trajectory: $B^k A^\omega$

$$\begin{pmatrix}1/3\\1/3\\1/3\end{pmatrix}=\begin{pmatrix}5/12\\1/3\\1/4\end{pmatrix}-\frac{1}{5}\begin{pmatrix}5/12\\0\\-5/12\end{pmatrix}$$

Trajectory: B^ω

When P_{init} converges towards $\begin{pmatrix}1/3\\1/3\\1/3\end{pmatrix}$

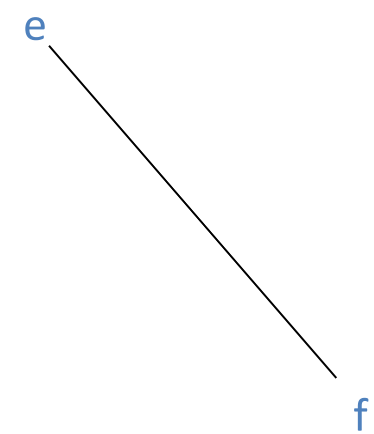
trajectory becomes $B^n A^\omega$ with n converging to ∞

=> Can show that language is $B^* A^\omega \cup B^\omega$

Not a finite union of trajectories

Language in general

Result: if all eigen values are distincts positive real numbers,
Then language is regular for Init a polytope.
[AGKV STACS'16]



e.g. : Set of Initial distributions:
$\{ \lambda e + (1-\lambda)f \mid \lambda \quad [0,1]\}$.

First, under these conditions, all trajectories are ultimately constant.

Ultimate Language

e A^ω

a_1(e)>0

$$\mathcal{L}_{ult}^{N_{max}}(H) = \{v \mid \exists w \in \{A,B\}^{N_{max}}, wv \in \mathcal{L}(H)\}$$

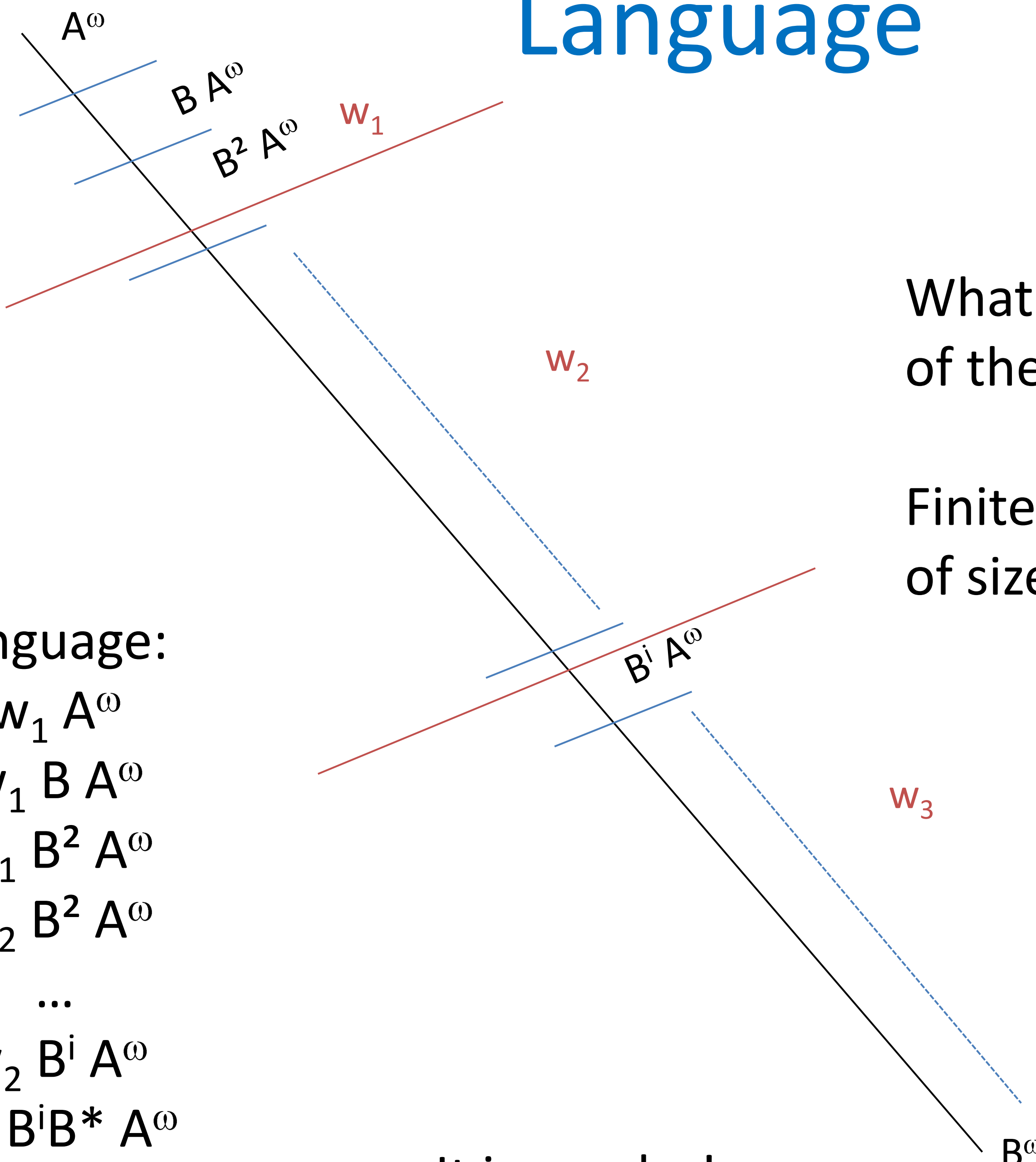
N_{max} such that after N_{max} steps,
the trajectories from e,f are A^ω and B^ω

The set of trajectories in (e,f) **after** N steps:
Lemma 1: Included into $B^* A^\omega$
Lemma 2: for all i, exists starting point with $B^i A^\omega$

a_1(f)=0
a_2(f)<0

f B^ω

Language



What about the prefixes of the N_{max} first steps?

Finite number of prefixes of size N_{max}.

Language:

$w_1\ A^{\omega}$

$w_1\ B\ A^{\omega}$

$w_1\ B^2\ A^{\omega}$

$w_2\ B^2\ A^{\omega}$

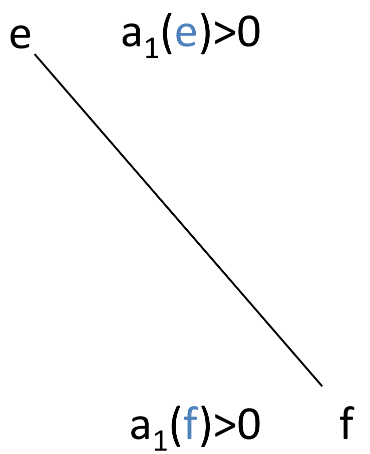
…

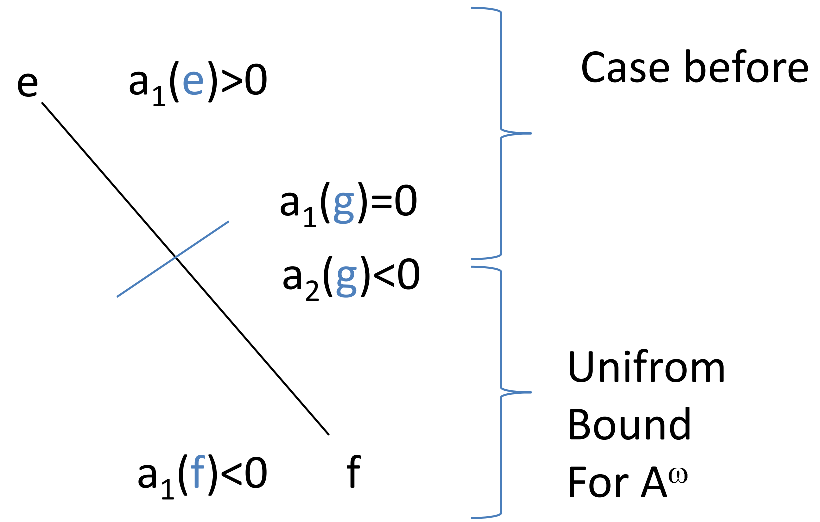
$w_2\ B^i\ A^{\omega}$

$w_3\ B^i B^*\ A^{\omega}$

It is regular!

In general with Polytope in 1D





max bound e, bound f
is a uniform bound
for utimately constant.

Polytopes in any Dimension

Case of $e_1..e_z$ extremities of Polytope with

$a_1(e_1) > 0$

$a_1(e_2)=0$, $a_2(e_2)<0$

$a_1(e_3)=a_2(e_3)=0$, $a_3(e_3)>0$

…

…. Sign $(a_k(e_k))=(-1)^k$

N is the max of the ultimately constant bound for $e_1..e_z$

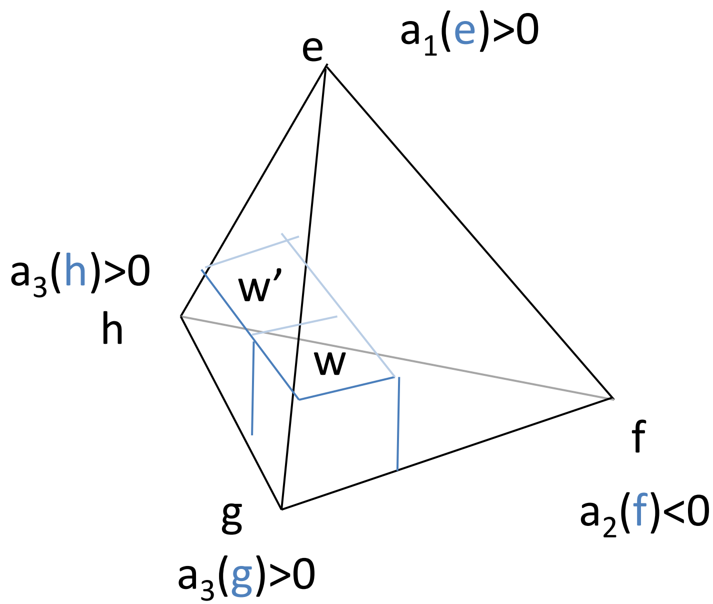
Result: $L^N_{ult}(H)= (A^*)\ B^*A^* \ldots B^*\ A^\omega$

The set of trajectories in $(e_1..e_z)$ **after** N steps:

Lemma 1: At most z switch, i.e. Included into $(A^*)\ B^*A^* \ldots B^*\ A^\omega$

Lemma 2: for all $i_1..i_z$, exists initial distrib with traj: $B^{i_1}\ A^{i_2} .. B^{i_z} A^\omega$

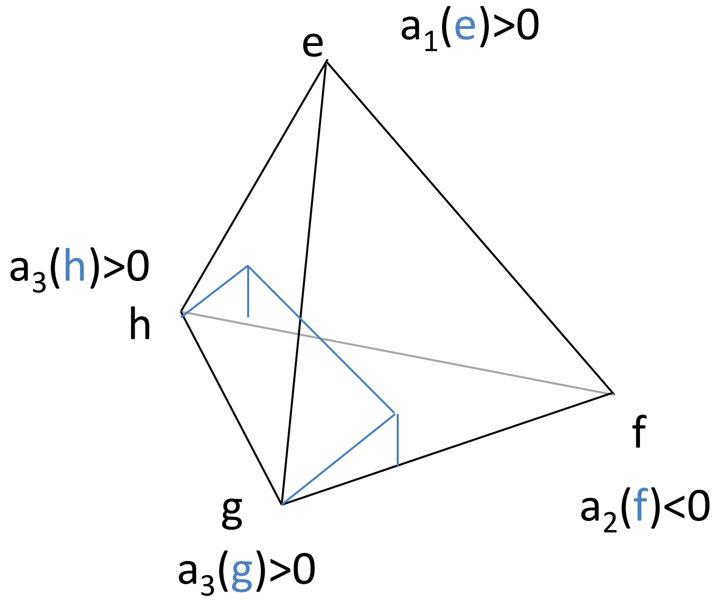
General Dimension: Handle Prefixes



Induction on the highest « z » in the space.

In the picture, z=3, n(dimension)=4
Take w touching (h,g) and touching (h,g,f) with a point not touching h or g
And touching (h,g,f,e) with a point not touching (hfg).
We can prove that for some i,
w A^i B A^ω is a trajectory continuity argument
w A^i B^ω is a trajectory => wA^iA*B*A^ω included into trajectory
w A^ω is a trajectory

In general



Induction on the highest « z » in the space.

Remove points with trajectory $wA^iA^*B^*A^\omega$ and $w'A^iA^*B^*A^\omega$
It remains a finite union of convex polyhedra with lower « z »

Hence the language is a finite union of regular set, hence it is regular.

Sum up

Property of eigenvalues of Markov chain	Regular language	Ultimately periodic traj.
Distinct, positive real numbers	✓ decidable	✓
Distinct, roots of real numbers	× Decidability ?	✓
Distinct	×	×

but approximable
[AAGT LICS'12]

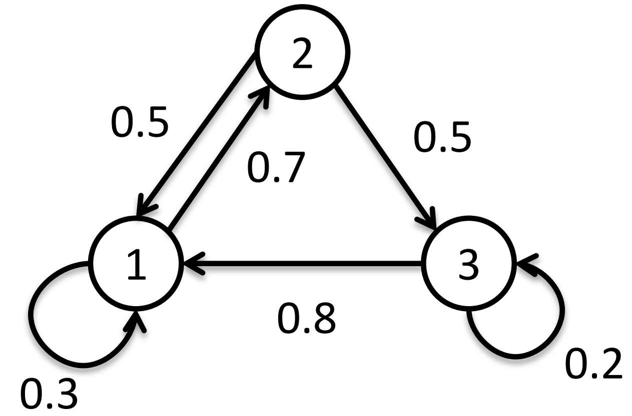
Distinct roots of real numbers: not regular.

$$M_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{512} & \frac{8r+3}{512} & \frac{3+3r}{64} & \frac{13+16r}{128} & \frac{9+2r}{32} & \frac{1+4r}{16} & \frac{1-r}{2} \end{bmatrix}$$

Approximation for Markov Chains.

Irreducible aperiodic chains

$$M = \begin{pmatrix} 0.3 & 0.7 & 0 \\ 0.5 & 0 & 0.5 \\ 0.8 & 0 & 0.2 \end{pmatrix}$$



M is irreducible aperiodic because:

$$M^2 = \begin{pmatrix} .44 & .21 & .35 \\ .55 & .35 & 0.1 \\ .4 & .56 & .04 \end{pmatrix}$$

Approximations for irreducible aperiodic chains:

Irreducible aperiodic: unique stationary distribution **f**.

Fix ε => exists K such that $|M^K \mathbf{u} - \mathbf{f}| < \varepsilon$ for all initial distribution **u**.

$A_1..A_n$ is an (ε,K)-approximate symbolic trajectory of
a concrete trajectory $d_1..d_n$ if
$d_i \in A_i$ for all i<K and d_i is ε-close to A_i for i>K.

Exact symbolic trajectory from init: ABBABBBABBBBA…

Epsilon => K=4,

Approx symbolic trajectories:

ABBABBAA…, ABBABBAB…, ABBABBBA…, ABBABBBB….

We get ABBABB (A or B)* is regular.

Approximations for irreducible aperiodic chains:

Th: Given MC + Init (set), it is decidable [AAGT, LICS'12] whether:

For some concrete trajectory w, there does not exists a approx trajectory satisfying ϕ,

=> w does not satisfies ϕ.

=> system does not satisfy ϕ.

For all concrete trajectory w, all approx traj satisfy ϕ

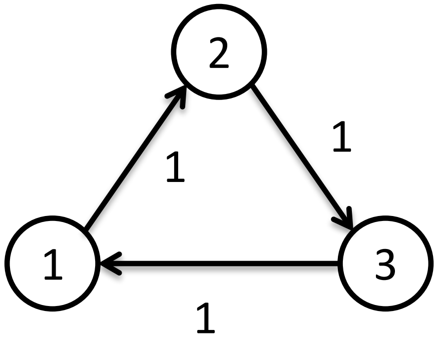
=> all w satisfies ϕ.

=> system satisfies ϕ.

Undetermined: for all concrete trajectory, there exists approximate trajectroty satisfying ϕ, but not for all.

=> Refine ε to reduce number of approx trajectories.

Irreducible Periodic chains

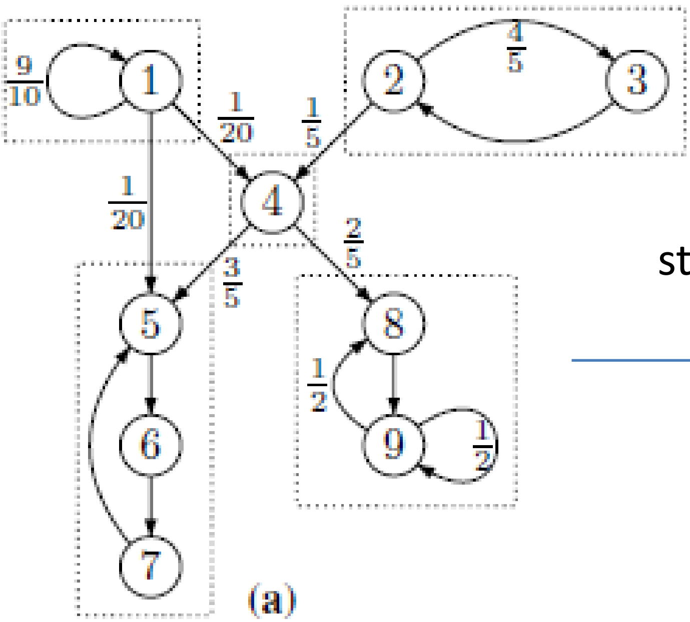


M is periodic of period 3.

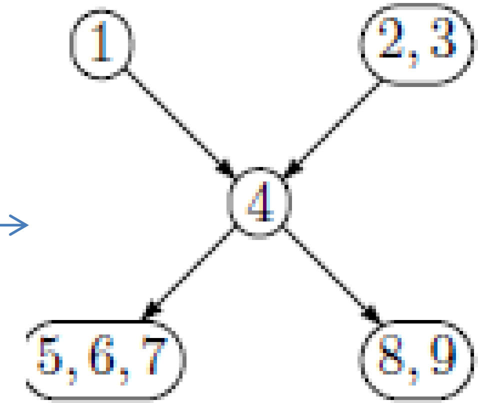
M^3 is irreducible aperiodic on disconnected partition of nodes.

Consider M^3 from Init,
Consider M^3 from M Init ,
Consider M^3 from M^2 Init

Not irreducible chains



Consider the strongly connected components.



Stationary distributions have weight 0 for non bottom SCC (1; 2-3, 4).
⇒Analyse the bottom SCC with earlier algorithm.

Tough part: Analyse non bottom SCC to get weights for bottom SCC, depending on Initial distribution (algorithm close to PCTL Mod. Check.)

+ uniform K over all initial distrib => allow to lift results to Languages

Polytope of initial Distributions

uniform K over all initial distrib => allow to lift results to Languages

Consider each extremities e1..en of the initial polytope.

Use linearity!

Compute the way they weight in the different BSCC

$e = \Sigma \lambda i\ ei$ => weight(e,BSCCk) = $\Sigma \lambda i$ weight(ei,BSCCk)

$\Rightarrow$Easy to compute the possible ultimately ε-reccuring set of letters

Then compute set of bounded prefixes with some ultimate set of letters, easy to compute as well.

Conclusion

Markov Chain (Unary PFA):
Simplistic formalism but still many open problems.

Even taking restrictive hypothesis,
not easy to describe their behavior.

But quantitative analysis of population is possible
under strong hypothesis or with approximations.