# MPRI

# Conservative approximation of polymers

**Jérôme Feret**
DI - ÉNS

http://www.di.ens.fr/~feret

Wednesday, the 23th of October, 2019

# On the menu today

1. Motivating example
2. Evolution systems
3. Box approximation
4. Symbolic reasoning
5. Conclusion

# An example with polymers

We denote by $A_n$ a chain of $n$ proteins.

We consider the following reactions (for $i, j \geq 1$):

1. $A_i + A_j \rightarrow A_{i+j} \qquad k$

2. $A_{i+j} \rightarrow A_i + A_j \qquad \begin{cases} k_d & \text{when } \; i = 1 \\ k_d + k'_d & \text{when } \; i > 1. \end{cases}$

# (Infinite) system of ODEs

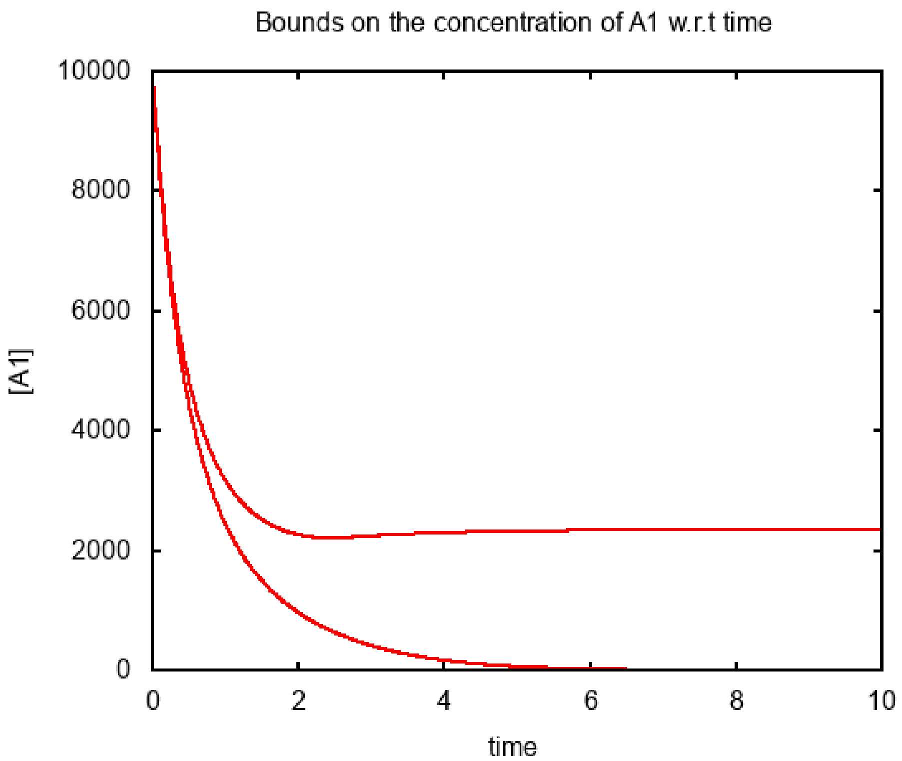
$$\frac{\mathrm{d}[A_n]}{\mathrm{d}t} = t_1^+(n) + t_2^+(n) + t_3^+(n) - t_1^-(n) - t_2^-(n) - t_3^-(n)$$

where:

$$\begin{aligned}
t_1^+(n) &\triangleq k \cdot \sum_{i+j=n} [A_i] \cdot [A_j]; \\
t_2^+(n) &\triangleq 2 \cdot k_d \cdot \sum_{i=n+1}^{+\infty} [A_i]; \\
t_3^+(n) &\triangleq \begin{cases} k'_d \cdot \sum_{i=3}^{+\infty} [A_i] & \text{if } n = 1, \\ k'_d \cdot \sum_{i=n}^{+\infty} ([A_{i+1}] + [A_{i+2}]) & \text{if } n \geq 2; \end{cases} \\
t_1^-(n) &\triangleq 2 \cdot k \cdot [A_n] \cdot \sum_{i=1}^{+\infty} [A_i]; \\
t_2^-(n) &\triangleq k_d \cdot (n-1) \cdot [A_n]; \\
t_3^-(n) &\triangleq \begin{cases} k'_d \cdot (n-2) \cdot [A_n] & \text{if } n \geq 3, \\ 0 & \textit{otherwise.} \end{cases}
\end{aligned}$$

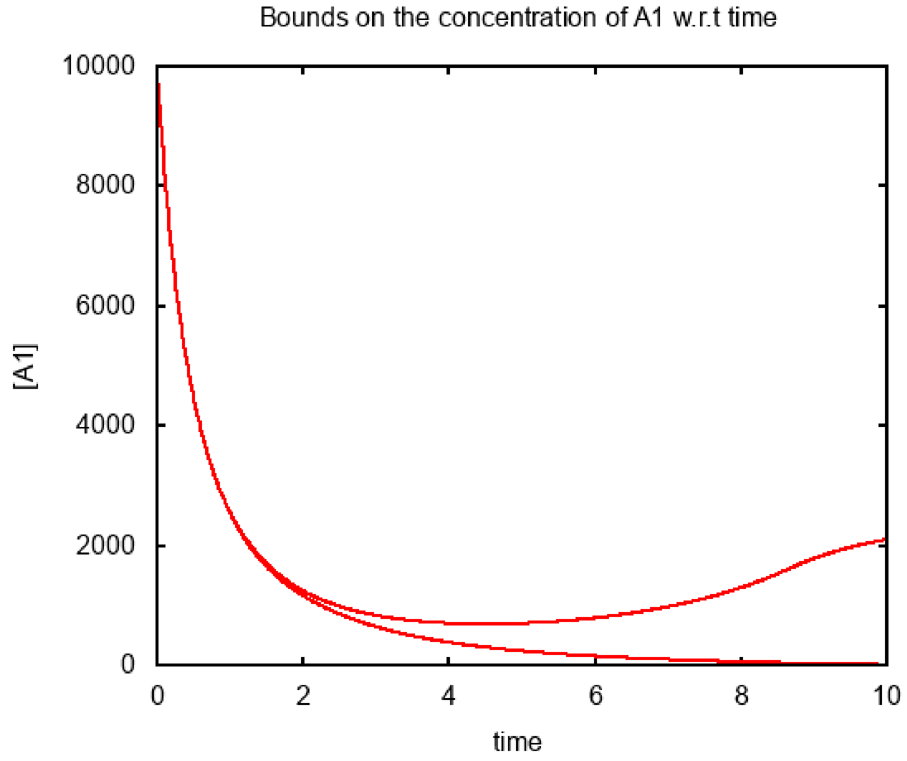
with the side condition: $\sum_{n \in \mathbb{N}} n \cdot [A_n] < +\infty$.

# Our goal:
# Bound the concentration of $A_1$



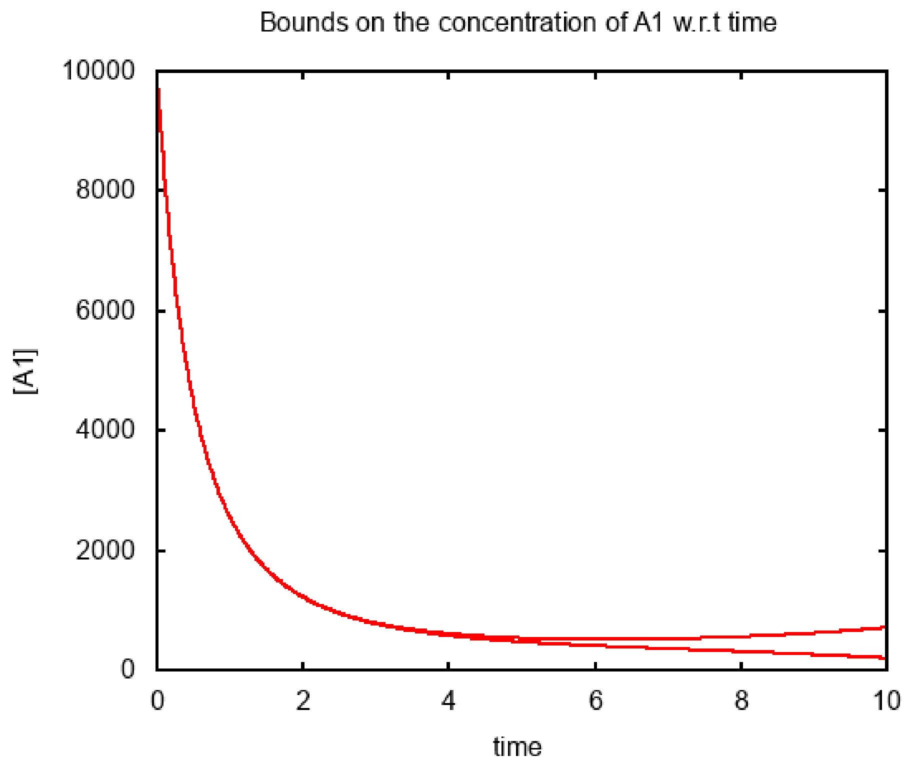

Obtained thanks to an ODEs of 18 variables.
(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Our goal:
# Bound the concentration of $A_1$



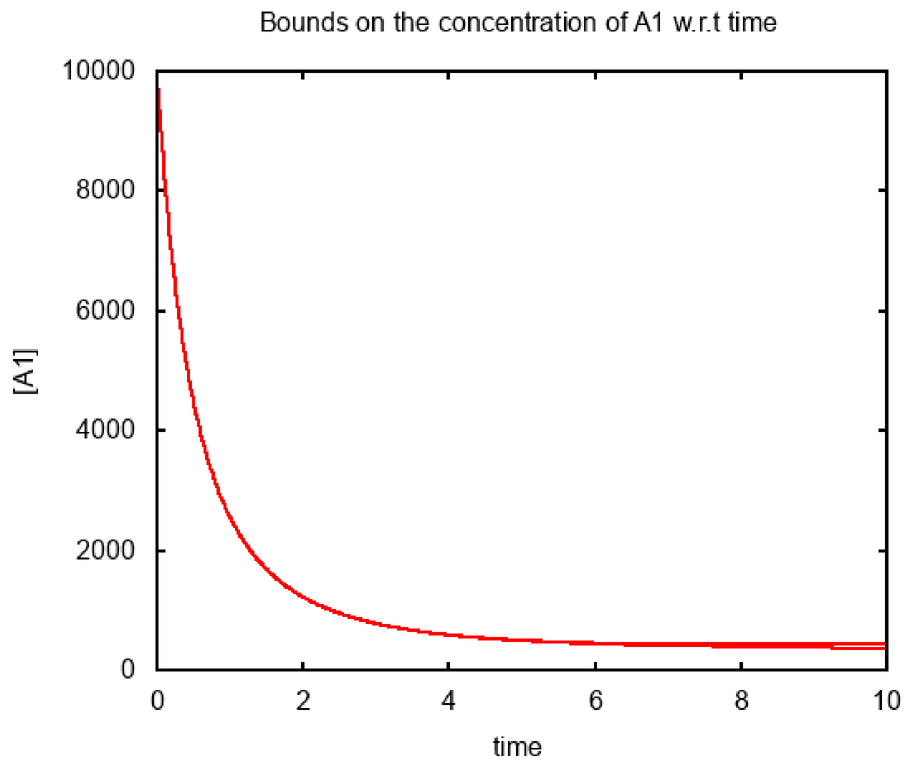

Obtained thanks to an ODEs of 36 variables.
(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Our goal:
# Bound the concentration of $A_1$



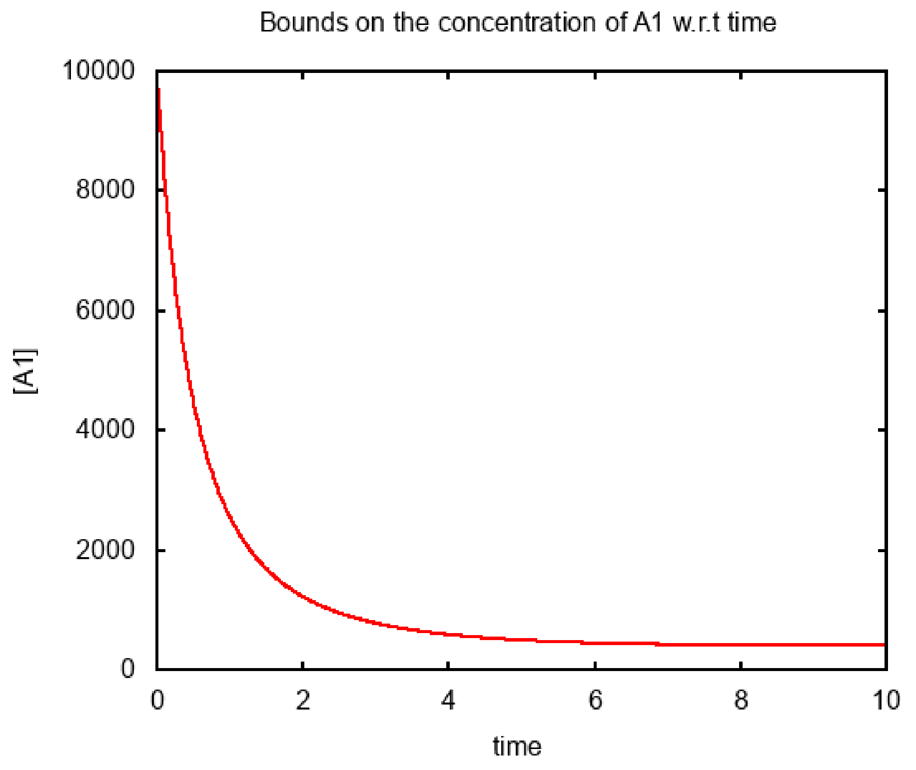

Obtained thanks to an ODEs of 54 variables.
(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Our goal:
# Bound the concentration of $A_1$




Obtained thanks to an ODEs of 72 variables.
(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Our goal:
# Bound the concentration of $A_1$




Obtained thanks to an ODEs of 90 variables.
(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Approach

1. Use a high level language to:
   (a) describe the model;
   (b) show the existence and unicity of the solution;
   (c) reason symbolically about some (potentially infinite) differentiable sums of variables:
      - express their derivatives,
      - infer inequalities among them;

2. Use box approximation to define a system of ODEs with two variables per sums of variables of interest (one for the lower bound, one for the upper bound) (error bounds are computed *a posteriori*).

# On the menu today

1. Motivating example
2. Evolution systems
3. Box approximation
4. Symbolic reasoning
5. Conclusion

# Definition

In a Banach space, a system defined as:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X,t) + G(X,t)$$

where:

1. $F$ is linear and induces a continuous semi-group (always satisfied when $F$ is triangular),
2. $G$ is Lipschitz on every bounded set;

is called an evolution system.

# Cade study

- Define the norm of a state as $\sum n \cdot |A_n|$.
- Define $F$ as the contribution of the reactions:

$$A_{i+j} \rightarrow A_i + A_j \qquad \begin{cases} k_d & \text{when } \; i = 1 \\ k_d + k'_d & \text{when } \; i > 1 \end{cases} \qquad \text{(for } i, j \geq 1\text{)}.$$

- Define $G$ as the contribution of the reactions:

$$A_i + A_j \rightarrow A_{i+j} \qquad k \qquad \text{(for } i, j \geq 1\text{)}.$$

# Properties

Evolutions systems:

- have exactely one maximal continuous solutions;
- maximal solutions are locally Lipschitz.
- whenever a maximal solution is not defined over $\mathbb{R}^+$, then the norm diverges;
- whenever $G$ is $\mathcal{C}_1$ and its derivative is bounded on bounded sets, then maximal solutions are also $\mathcal{C}_1$.

[Hundertmark et al., Operator Semigroups and Dispersive Equations, 16th Internet Seminar on Evolution Equations 2013]

# On the menu today

1. Motivating example
2. Evolution systems
3. Box approximation
4. Symbolic reasoning
5. Conclusion

# Principle

Given a finite system of ODEs:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X,t).$$

Box approximation:

1. approximates the state of the system by a (hyper)-box
   (twice many variables as in the initial system)
2. associates to each (hyper)-face an expression that bounds conservatively the partial derivative of the system with respect to the corresponding variable over this (hyper)-face.

Sound whenever $F$ is locally Lipschitz w.r.t to the state and continuous w.r.t time.

[M. Kirkilionis and S. Walcher, On comparison systems for ordinary differential equations, J. Math. Anal. Appl. 299 (2004)]

# Example: ODEs

Consider the following system of ODEs:

$$\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = y \cdot (2 - \cos(y)) - x \cdot (2 - \sin(y)) \\ \frac{\mathrm{d}y}{\mathrm{d}t} = x \cdot (2 - \cos(y)) - y \cdot (2 - \sin(y)) \\ x(0) = y(0) = 1 \end{cases}$$

# Example: Invariants

Consider the following system of ODEs:

$$\begin{cases} \frac{dx}{dt} = y \cdot (2 - \cos(y)) - x \cdot (2 - \sin(y)) \\ \frac{dy}{dt} = x \cdot (2 - \cos(y)) - y \cdot (2 - \sin(y)) \\ x(0) = y(0) = 1 \end{cases}$$

We have:

$$\begin{cases} y - 3 \cdot x \leq \frac{dx}{dt} \leq 3 \cdot y - x \\ x - 3 \cdot y \leq \frac{dy}{dt} \leq 3 \cdot x - y. \end{cases}$$

# Example: Box approximation

Thus, the following system of ODEs:

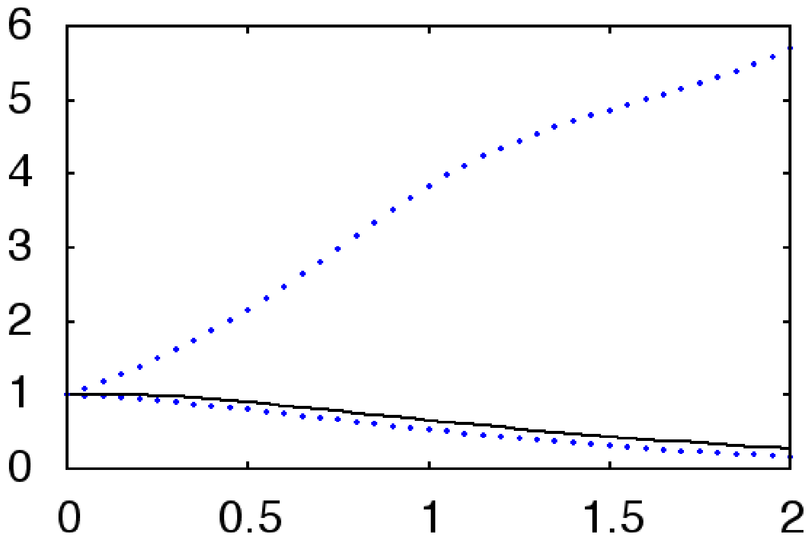
$$\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = y \cdot (2 - \cos(y)) - x \cdot (2 - \sin(y)) \\ \frac{\mathrm{d}y}{\mathrm{d}t} = x \cdot (2 - \cos(y)) - y \cdot (2 - \sin(y)) \\ x(0) = y(0) = 1 \end{cases}$$

can be safely approximated by the following one:

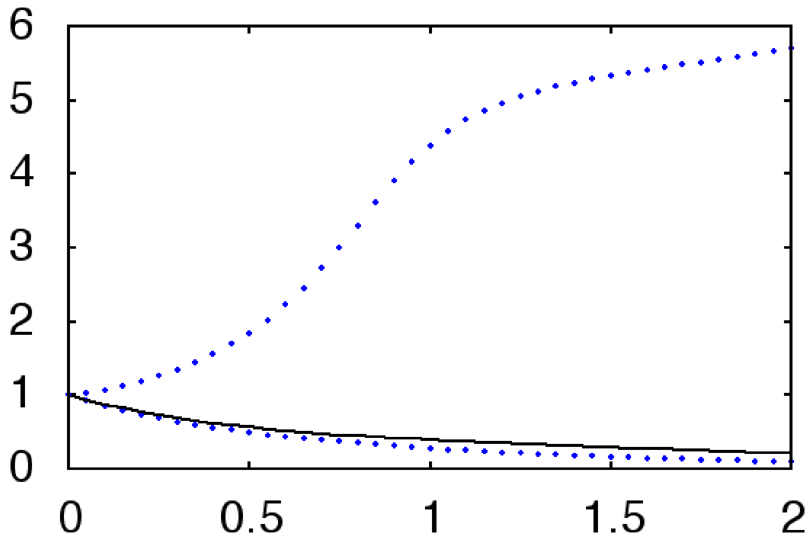
$$\begin{cases} \frac{\mathrm{d}\underline{x}}{\mathrm{d}t} = \underline{y} - 3 \cdot \underline{x} \\ \frac{\mathrm{d}\overline{x}}{\mathrm{d}t} = 3 \cdot \overline{y} - \overline{x} \\ \frac{\mathrm{d}\underline{y}}{\mathrm{d}t} = \underline{x} - 3 \cdot \underline{y} \\ \frac{\mathrm{d}\overline{y}}{\mathrm{d}t} = 3 \cdot \overline{x} - \overline{y} \\ \underline{x} = \overline{x} = \underline{y} = \overline{y} = 1 \end{cases}$$

# Example: Numerical results

x(t) with respect to t

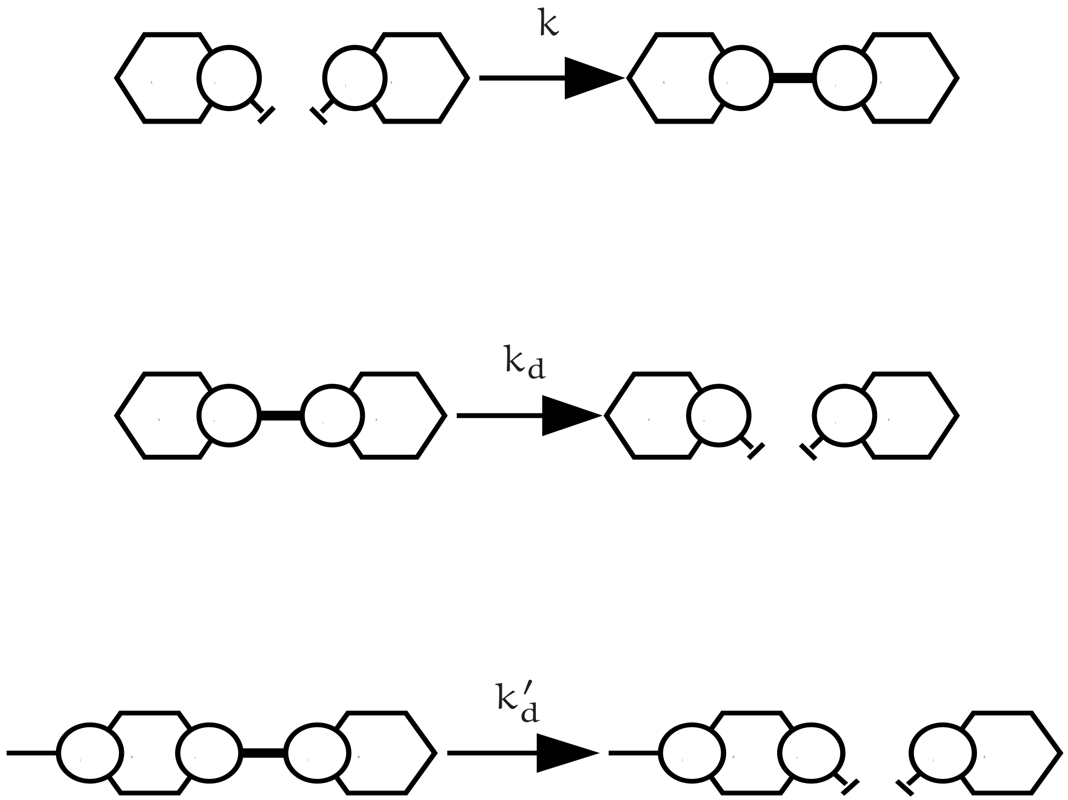



y(t) with respect to t



# On the menu today

# The model in Kappa



# Back to evolution systems

Let us consider both following definitions:

1. A rule is dispersive if it is unary and it splits its pattern into smaller ones.
2. A rule is locally Lipschitz if for every pattern $P$ the number of embedding from the patterns in the lhs of the rules and the pattern $P$ is uniformly bounded.

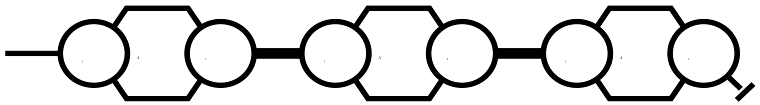
Then: A finite set of Kappa rules such that every rule is either dispersive, or locally Lipschitz (or both) induces an evolution system, for the norm defined as the overall concentration of proteins.

# Connected patterns

Connected patterns have both an intensional and an extensional meaning.

A connected pattern may be seen:

1. as a connected graph:



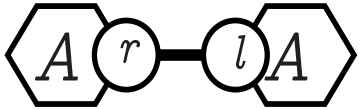
2. as a linear (potentially infinite) sum of fully specified connected graphs:

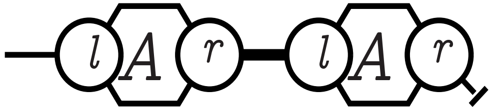
$$[-A_3\dashv] = \sum_{n\geq 4} [\vdash A_n \dashv]$$

# List of connected patterns

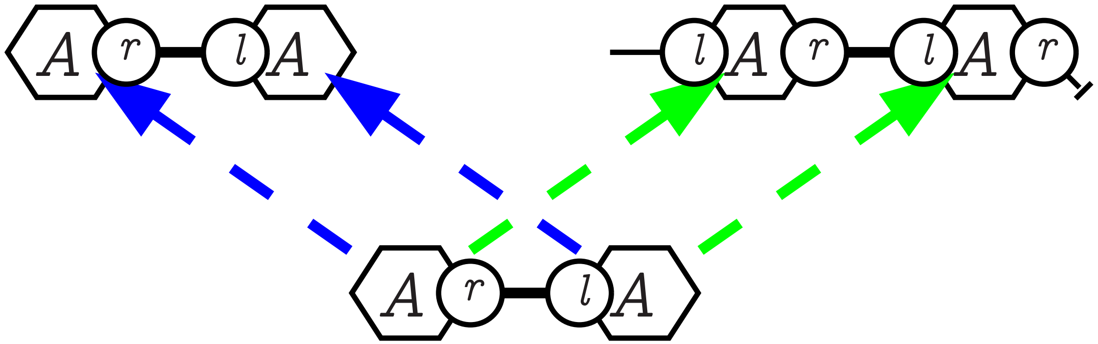
- $[\vdash A_n \dashv]$: concentration of polymer of length $n$;
- $[\vdash A_n]$: concentration of polymer of length at least $n$;
- $[A_n \dashv]$: concentration of polymer of length at least $n$;
- $[\vdash A_n -]$: concentration of polymer of length at least $n+1$;
- $[- A_n \dashv]$: concentration of polymer of length at least $n+1$;
- $[A_n] = \sum_{i \in \mathbb{N}} (i+1) \cdot [\vdash A_{n+i} \dashv]$;
- $[A_n -] = \sum_{i \in \mathbb{N}} (i+1) \cdot [\vdash A_{n+1+i} \dashv]$;
- $[- A_n] = \sum_{i \in \mathbb{N}} (i+1) \cdot [\vdash A_{n+1+i} \dashv]$;
- $[- A_n -] = \sum_{i \in \mathbb{N}} (i+1) \cdot [\vdash A_{n+2+i} \dashv]$.

# Proper overlap between connected patterns

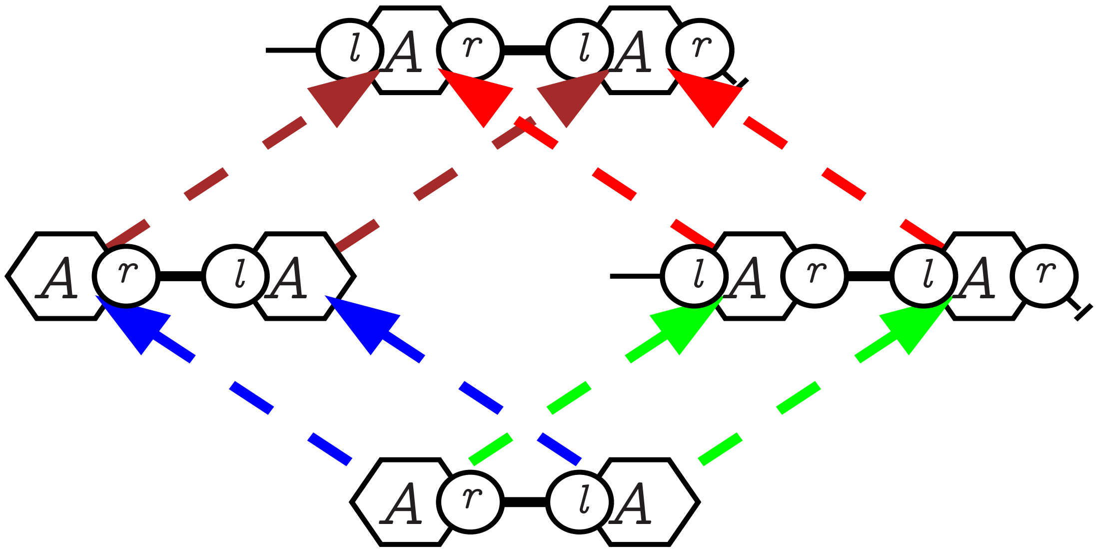


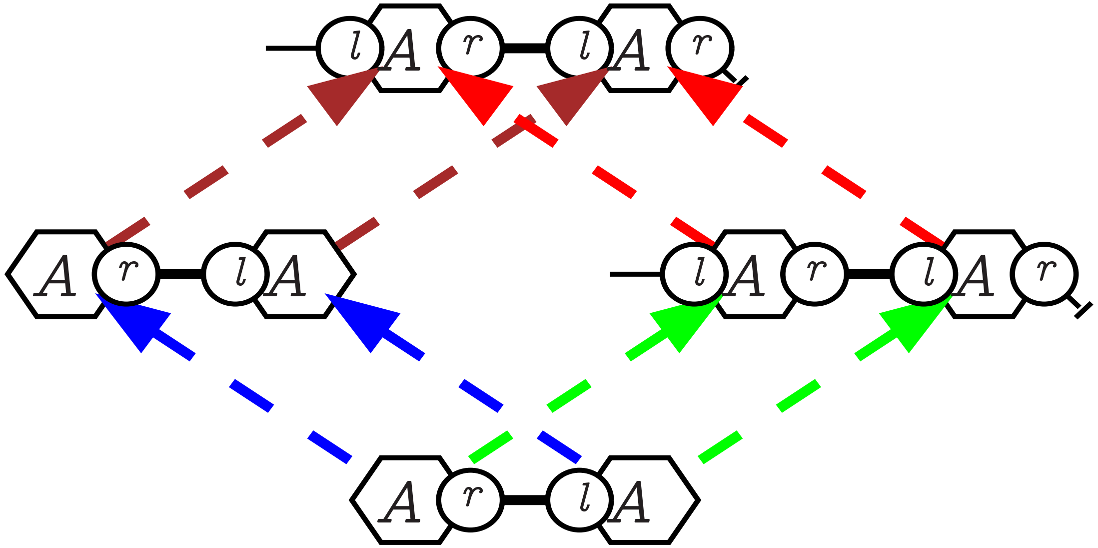
# Proper overlap between connected patterns



# Proper overlap between connected patterns



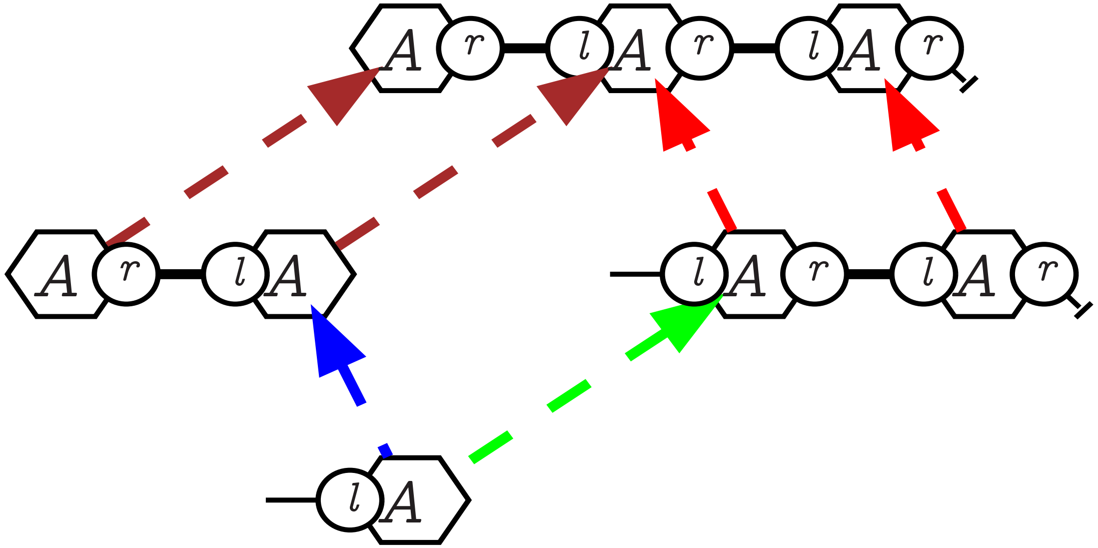
# Proper overlap between connected patterns



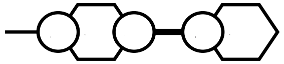
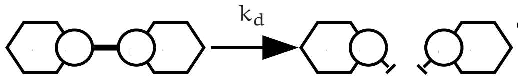

To count properly, the common partern shall be taken maximal (pullback). The define the gluing of two patterns, the unifying partern shall be taken minimal (pushout).

# Proper overlap between connected patterns

There may be several overlaps between two patterns.

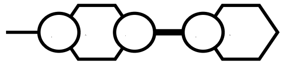
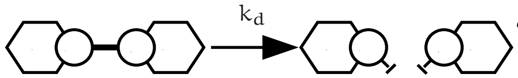


# Proper consumption

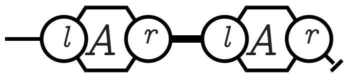
Which quantity of the pattern 
is consumed due to the rule 

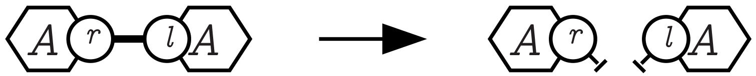
?

# Proper consumption

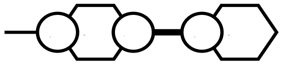
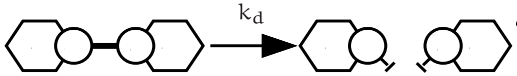
Which quantity of the pattern 
is consumed due to the rule 

?

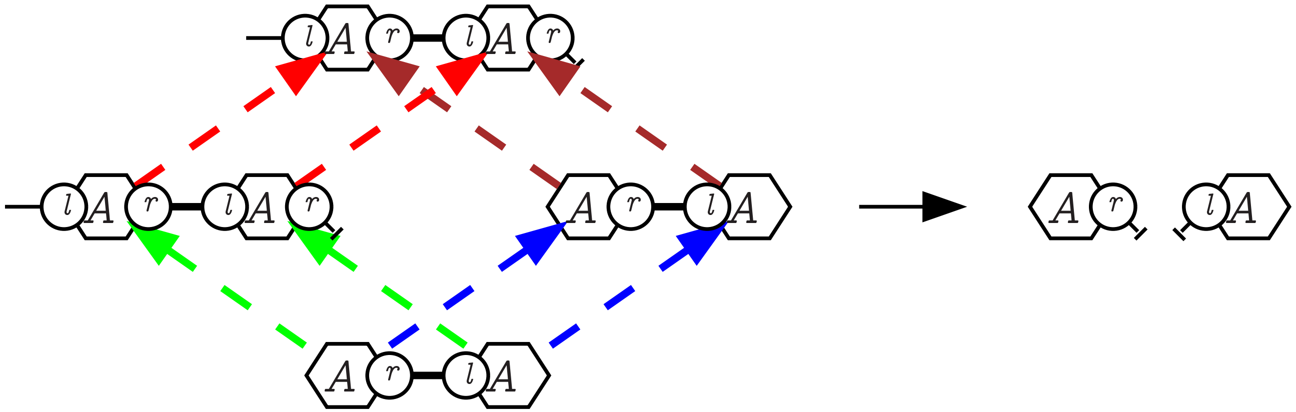


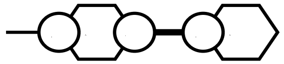
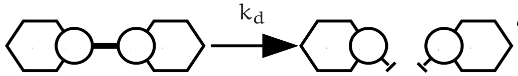
# Proper consumption

Which quantity of the pattern 
is consumed due to the rule 

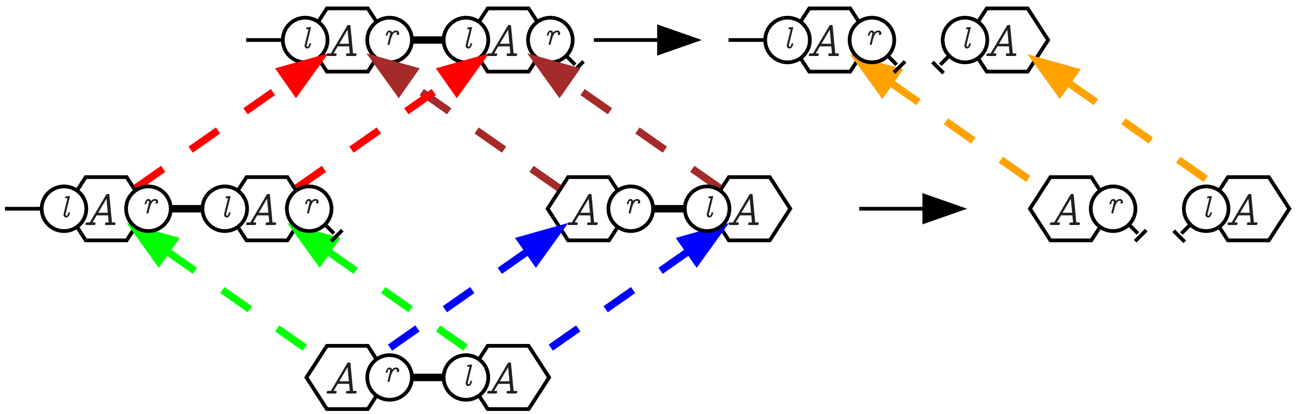
?



# Proper consumption

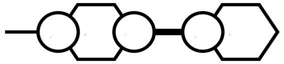
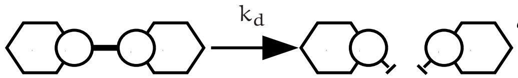
Which quantity of the pattern 
is consumed due to the rule 

?

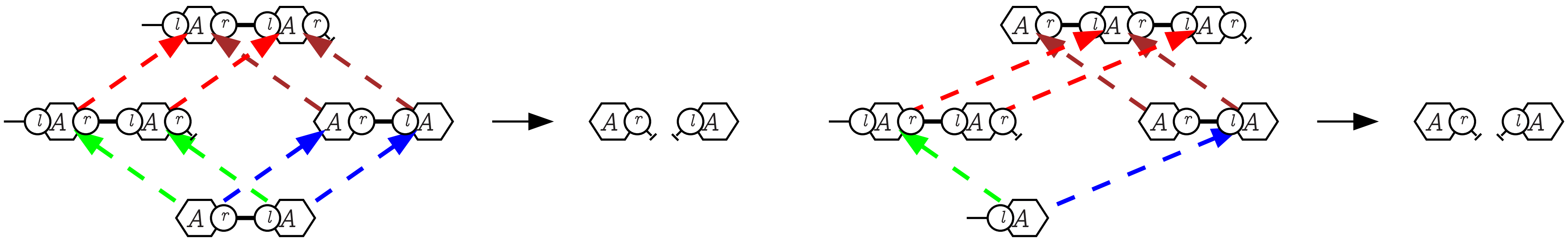



$$k_d \cdot [-A_2\dashv].$$

# Proper consumption

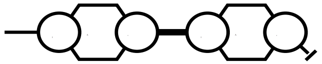
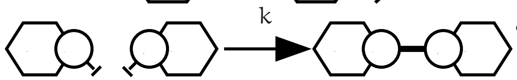
Which quantity of the pattern 
is consumed due to the rule 

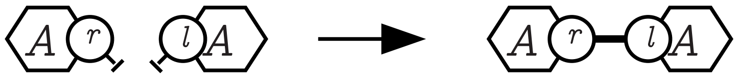
?

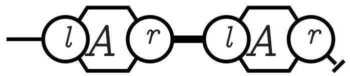


$$k_d \cdot ([-A_2\dashv] + [A_3\dashv]).$$

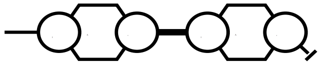
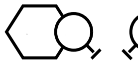
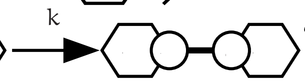
# Proper production

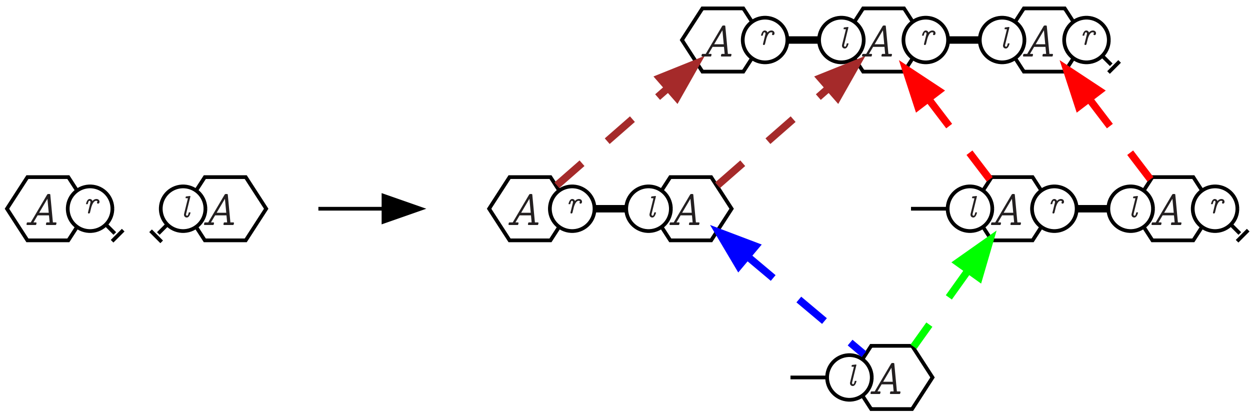
Which quantity of the pattern 
is produced due to the rule ?



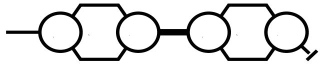



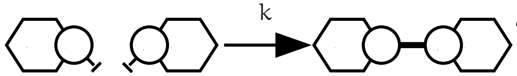
# Proper production

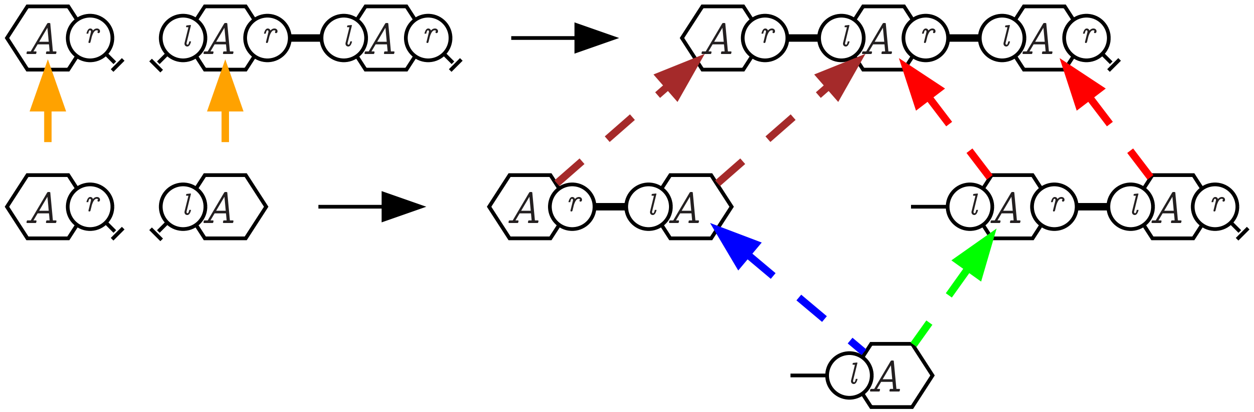
Which quantity of the pattern 
is produced due to the rule  ?



# Proper production

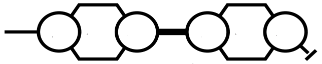
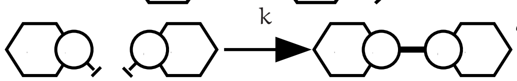
Which quantity of the pattern 

is produced due to the rule ?

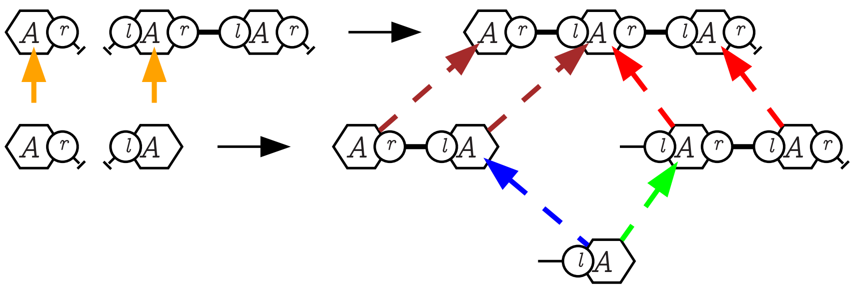


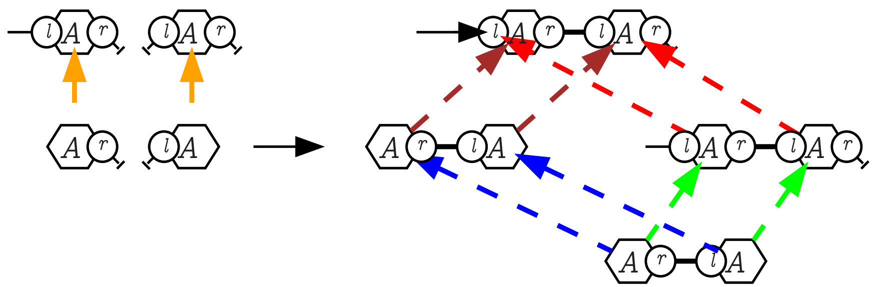
$$k \cdot [A_1 \dashv] \cdot [\vdash A_2 \dashv].$$

# Proper production

Which quantity of the pattern 
is produced due to the rule 

?





$$k \cdot ([A_1 \dashv] \cdot [\vdash A_2 \dashv] + [-A_1 \dashv] \cdot [\vdash A_1 \dashv]).$$

# Exact derivatives

$$\frac{\mathrm{d}[\vdash A_n \dashv]}{\mathrm{d}t} = \mathcal{T}^+_{1,1}(n) + \mathcal{T}^+_{1,2}(n) + \mathcal{T}^+_{1,3}(n) - \mathcal{T}^-_{1,1}(n) - \mathcal{T}^-_{1,2}(n) - \mathcal{T}^-_{1,3}(n)$$

where:

$$\mathcal{T}^+_{1,1}(n) \triangleq k \cdot \sum_{i+j=n} [\vdash A_i \dashv] \cdot [\vdash A_j \dashv];$$

$$\mathcal{T}^+_{1,2}(n) \triangleq k_d \cdot ([\vdash A_{n+1}] + [A_{n+1} \dashv]);$$

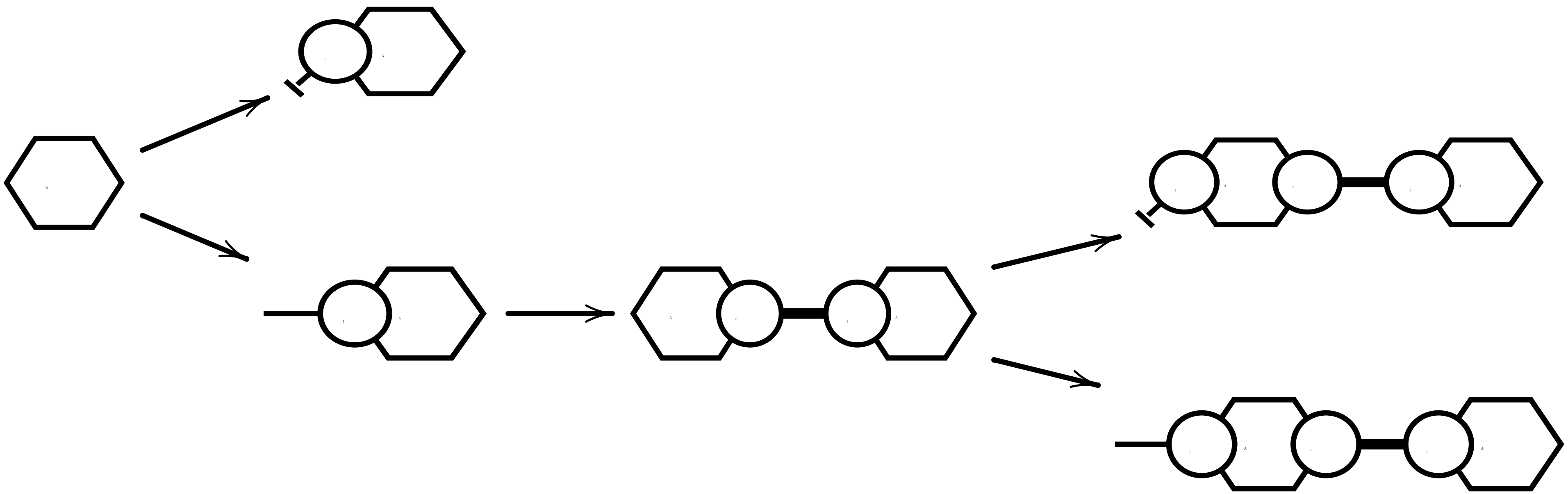
$$\mathcal{T}^+_{1,3}(n) \triangleq \begin{cases} k'_d \cdot [- A_{n+1} \dashv] & \text{if } n = 1 \\ k'_d \cdot ([- A_{n+1} \dashv] + [\vdash A_{n+1}]) & \text{if } n \geq 2; \end{cases}$$

$$\mathcal{T}^-_{1,1}(n) \triangleq k \cdot [\vdash A_n \dashv] \cdot ([\vdash A_1] + [A_1 \dashv]);$$

$$\mathcal{T}^-_{1,2}(n) \triangleq k_d \cdot (n-1) \cdot [\vdash A_n \dashv];$$

$$\mathcal{T}^-_{1,3}(n) \triangleq \begin{cases} k'_d \cdot (n-2) \cdot [\vdash A_n \dashv] & \text{if } n \geq 3 \\ 0 & \text{otherwise.} \end{cases}$$

# Orthogonal refinement



# Inequalities

1. $[\vdash A_n \dashv] \leq \dfrac{[A_1] - \sum_{k=1,k\neq n}^{N} k \cdot [\vdash A_k \dashv]}{n}$;
2. $[A_n] \leq [A_1] - \sum_{k=1}^{n-1} k \cdot [\vdash A_k \dashv]$;
3. $[A_n -] \leq [A_1] - \sum_{k=1}^{n-1} k \cdot [\vdash A_k \dashv]$;
4. $[\diamond_l A_n \dashv] \leq \dfrac{[A_1] - \sum_{k=1}^{n-1} k \cdot [\vdash A_k \dashv]}{n-1} \qquad \forall \diamond_l \in \{\vdash, -, \varepsilon\}$;
5. $[\vdash A_n \diamond_r] \leq \dfrac{[A_1] - \sum_{k=1}^{n-1} k \cdot [\vdash A_k \dashv]}{n-1} \qquad \forall \diamond_r \in \{\dashv, -, \varepsilon\}$;.

# Exact derivatives

$$\frac{\mathrm{d}[\vdash A_1 -]}{\mathrm{d}t} = \mathcal{T}^{+}_{2,1} + \mathcal{T}^{+}_{2,2} + \mathcal{T}^{+}_{2,3} - \mathcal{T}^{-}_{2,1} - \mathcal{T}^{-}_{2,2} - \mathcal{T}^{-}_{2,3}$$

where:

$$\begin{aligned}
\mathcal{T}^{+}_{2,1} &\triangleq k \cdot ([\vdash A_1 \dashv] \cdot [\vdash A_1]); \\
\mathcal{T}^{+}_{2,2} &\triangleq k_d \cdot [A_2 -]; \\
\mathcal{T}^{+}_{2,3} &\triangleq k'_d \cdot [- A_2 -]; \\
\mathcal{T}^{-}_{2,1} &\triangleq k \cdot [\vdash A_1 -] \cdot [A_1 \dashv]; \\
\mathcal{T}^{-}_{2,2} &\triangleq k_d \cdot [\vdash A_2]; \\
\mathcal{T}^{-}_{2,3} &\triangleq 0.
\end{aligned}$$

# The ODEs

$$\frac{\mathrm{d}\underline{[\vdash A_1 -]}}{\mathrm{d}t} = \underline{t^{+}_{2,1}} + \underline{t^{+}_{2,2}} + \underline{t^{+}_{2,3}} - \overline{t^{-}_{2,1}} - \overline{t^{-}_{2,2}} - \overline{t^{-}_{2,3}},$$

where:

$$\underline{t^{+}_{2,1}} \triangleq k \cdot max(0, \underline{[\vdash A_1 \dashv]}) \cdot (max(0, \underline{[\vdash A_1 \dashv]}) + max(0, \underline{[\vdash A_1 -]}));$$

$$\underline{t^{+}_{2,2}} \triangleq k_d \cdot max(0, \underline{[- A_1 -]});$$

$$\underline{t^{+}_{2,3}} \triangleq k'_d \cdot max(0, \underline{[- A_2 -]});$$

$$\overline{t^{-}_{2,1}} \triangleq k \cdot max(0, \underline{[\vdash A_1 -]}) \cdot \left( min \left( \underline{[\vdash A_1 -]}, \overline{[A_1]} - \sum_{n=2}^{N} n \cdot \underline{[\vdash A_n \dashv]} \right) + min \left( \overline{[- A_1 \dashv]}, \overline{[A_1]} - \sum_{n=2}^{N} n \cdot \underline{[\vdash A_n \dashv]} \right) \right);$$

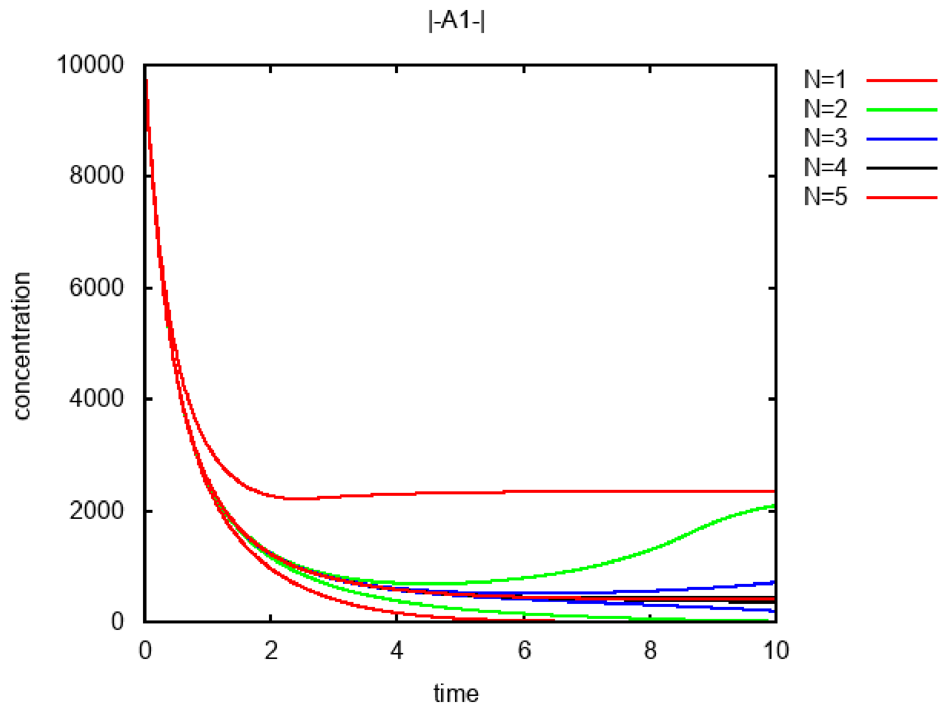
$$\overline{t^{-}_{2,2}} \triangleq k_d \cdot max(0, \underline{[\vdash A_1 -]});$$

$$\overline{t^{-}_{2,3}} \triangleq 0.$$

# On the menu today

1. Motivating example
2. Evolution systems
3. Box approximation
4. Symbolic reasoning
5. Conclusion

# Numerical results




(with parameters $[A_1]_0 = 10000$, $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$).

# Conclusion

- We can deal models of polymers that:
  1. are finitely branching;
  2. and locally defined.
- High-level languages enable to denote some infinite sums of variables and to handle them symbolically

  (Prove that they are differentiable, express their derivative, compare them).
- Box approximation can be used to derive time-dependent bounds on the values of some observables

  (Safe numerical bounds are computed *a posteriori* )
  (Approximation locally adapts to the state of the system)
  (Partial derivatives are considered only on the corresponding (hyper)-faces).