Conservative approximations of polymers

Ken Chanseau Saint-Germain[1] and Jérôme Feret[1][0000−0002−2601−2785]

DI-ENS (INRIA/ENS/CNRS),
PSL Research University,
75005 Paris, France
{chanseau,feret}@ens.fr

Abstract. We propose a systematic approach to approximate the behaviour of models of polymers synthesis/degradation. Our technique consists in discovering time-dependent lower and upper bounds for the concentration of some patterns of interest. These bounds are obtained by approximating the state of the system by a hyper-box, with differential equations defining the evolution of the coordinates of each hyper-face. The equation of each hyper-face is obtained by pessimistically bounding the derivative with respect to the corresponding coordinate when the system state ranges over this hyper-face.
In order to synthesise these bounds, we use Kappa to describe our models of polymers. This provides symbolic equalities and inequalities which intentionally may be understood as algebraic constructions over patterns, and extensionally as sound properties about the concentration of the biomolecular species that contain these patterns.

Keywords: Polymers · Differential equations · Evolution systems · Conservative approximation · Rule-based modelling.

1 A model of polymers synthesis/degradation

We study here a simple system of polymers with only one kind of protein having two identified sites that we will call left and right. We assume that the rightmost site of a polymer instance may bind to the leftmost site of another polymer instance at rate k. We also assume that the leftmost bond in a polymer is stronger that the other bonds. The leftmost bond may break at rate k_d whereas the other bonds may be broken at rate $k_d + k'_d$.

We denote as $[\vdash A_n \dashv]$ the concentration of polymers made of n connected proteins. The semantics of our system is obtained by applying the principle of Mass action. More precisely, the following system of equations:

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

where:

– $t_1^+(n) \triangleq k \cdot \sum_{i+j=n} [\vdash A_i \dashv] \cdot [\vdash A_j \dashv]$; $t_2^+(n) \triangleq 2 \cdot k_d \cdot \sum_{i=n+1}^{+\infty} [\vdash A_i \dashv]$;

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

with the side-condition that the infinite sum $\sum_{n \in \mathbb{N}} n \cdot [\vdash A_n \dashv]$ is converging, is an evolution system in the Banach space of the tempered sequences of real numbers (with the norm $\sum_{n \in \mathbb{N}} n \cdot [\vdash A_n \dashv]$). Indeed, this system is made of a linear part inducing a continuous semi-group and a second part that is Lipschitz on every bounded subset. As a consequence [3, Th 8.6], it has a unique maximal continuous solution. Since additionally, the system is norm invariant, the maximal solution is defined over the interval $[0, +\infty)$ [3, Th 8.6].

2 Box approximation

Theory of differential equation comparison allows to bound the state of finite ordinary differential systems by a time-dependent hyper-box. For this we need one function to continuously update the position of each hyper-face. A given coordinate is associated with a lower and an upper hyper-face. Whenever, for each coordinate, the function associated to the upper hyper-face over-estimates over the whole upper hyper-face the derivative of the initial function with respect to this coordinate and the function associated to the lower hyper-face under-estimates over the whole lower hyper-face the same derivative, then the solution of the initial system remains in the time-dependent hyper-box (that is defined as the solution of the approximate equations) [4, Prop 1.7]. This proposition requires all the functions to be continuous with respect to time and locally Lipschitz with respect to the state of the system.

3 Kappa

The models that are generated by a finite number of generic context-free mechanisms, can be described in Kappa [2]. Not only, this provides a finite description of them by the means of rewriting rules, but also patterns are good candidates to define observables of interest. Intensionally, a pattern P is a part of a bio-molecular species. Extensionally it may be interpreted as a function $[P]$ mapping each Kappa graph G to the set of the embeddings from P to G, or even as a function $[\![P]\!]$ from the states of the system to the sum of the concentrations of each bio-molecular species containing occurrences of the pattern P.

Kappa comes with universal categorical constructions to reason symbolically over the extensional interpretation of patterns. Their soundness is formally proven once for all at the language level. In particular, we use two kinds of construction. *Orthogonal refinements* (e.g see Fig. 1) refine patterns by the means of decision trees by gradually inserting new information about the state of sites.

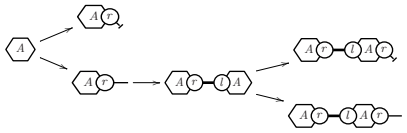

Fig. 1. An orthogonal refinement.

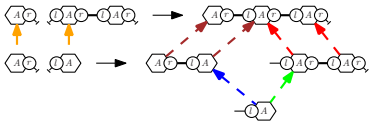

Fig. 2. Specialisation of the rule that binds two polymers to the case where the second agent is the first agent of a pattern made of two connected proteins, the first one with a left site bound and the second one the second site free.

Each width-cut of such decision tree induces a partition over the set of the embeddings that is denoted by the pattern at the root of the tree, hence ensuring preservation of the concentration of the initial pattern. *Rule refinements* (e.g see Fig. 2) specialise rules for the consumption or the production of a given pattern according to a potential overlap between this pattern and the left hand side or the right hand side of a rule. From an algebraic point of view, rule refinements are defined by the means of pushouts in the category of partial embeddings. Interestingly, these constructions allow to prove that the concentration of each pattern is differentiable and to express the derivative of each pattern as an expression of the concentration of some patterns. This result is proven in [1, Chp 5] in the finite case. Extension to infinite systems requires the permutation of two limits, which has been proven once for all at the language level.

Only a subset of Kappa induces evolution systems. In particular we require that every rule is either a unary one that splits patterns into smaller ones, or a rule where each connected component in the left hand side has a free site. We also require that the number of free sites in reachable bio-molecular species is uniformly bounded. By rigidity [5], such sets of rules induce evolution systems.

4 Application

We use our framework to bound the concentration of monomers in our model. Our abstraction is parameterised by a number N. We keep as patterns the $9 \cdot N$ (according to their length and whether the leftmost and the rightmost sites are free, bound, or not specified) patterns containing at most N connected agents. There are two variables per patterns. In Fig. 3, we show the bounds that we obtain with $N = 3$, the rate constants $k = 10^{-4}$, $k_d = 10^{-2}$, and $k'_d = 10^{-1}$, and an initial state made of monomers only at concentration 10000. By construction, the curves provide time-dependent lower and upper bounds on the concentration of monomers in the initial model.

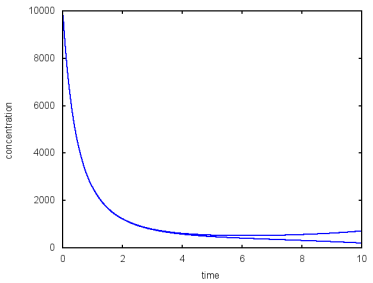

Fig. 3: Time-dependent bounds on the concentration of monomers.

References

1. Camporesi, F.: Réduction formelle et exacte de modèles de voies de signalisation en Kappa (Exact and formal reduction of models of signaling pathways written in Kappa). Ph.D. thesis, École normale supérieure, Paris, France (2017)
2. Danos, V., Laneve, C.: Formal molecular biology. Theor. Comput. Sci. **325**(1), 69–110 (2004). https://doi.org/10.1016/j.tcs.2004.03.065, https://doi.org/10.1016/j.tcs.2004.03.065
3. Hundertmark, D., Meyries, M., Machinek, L., Schnaubelt, R.: Operator semigroups and dispersive equations. In: Lecture Notes of the 16th Internet Seminar on Evolution Equations (2013), http://www.math.kit.edu/iana3/ schnaubelt/media/isem16-skript.pdf
4. Kirkilionis, M., Walcher, S.: On comparison systems for ordinary differential equations. Journal of Mathematical Analysis and Applications **299**(1), 157 – 173 (2004)
5. Petrov, T., Feret, J., Koeppl, H.: Reconstructing species-based dynamics from reduced stochastic rule-based models. In: Laroque, C., Himmelspach, J., Pasupathy, R., Rose, O., Uhrmacher, A.M. (eds.) Winter Simulation Conference. WSC (2012)
6. Saint-Germain, K.C.: Numerically approximated model reduction (2016), Master thesis, Parisian Research Master in Computer Science (MPRI))