Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A proof theoretic view of spatial and temporal dependencies in biochemical systems

C. Olarte [a], D. Chiarugi [b], M. Falaschi [c,*], D. Hermith [c]

[a] *ECT, Universidade Federal do Rio Grande do Norte, Brazil*
[b] *Max Planck Institute of Colloids and Interfaces, Potsdam-Golm, Germany*
[c] *Dept. of Information Engineering and Mathematics, Via Roma 56, Siena, Italy*

A R T I C L E   I N F O

A B S T R A C T

The behavior of biochemical systems such as metabolic and signaling pathways may depend on either the location of the reactants or on the time needed for a reaction to occur. In this paper we propose a formalism for specifying and verifying properties of biochemical systems that combines, coherently, temporal and spatial modalities. To this aim, we consider a fragment of intuitionistic linear logic with subexponentials (SELL). The subexponential signature allows us to capture the spatial relations among the different components of the system and the timed constraints. We illustrate our approach by specifying some well-known biological systems and verifying properties of them. Moreover, we show that our framework is general enough to give a logic-based semantics to P systems. We show that the proposed logical characterizations have a strong level of adequacy. Hence, derivations in SELL follow exactly the behavior of the modeled system.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

In the last decade, the impressive enhancement of experimental techniques in biology has made available a huge amount of information concerning living organisms. In this way, the knowledge about the components of biological systems is becoming increasingly detailed and accurate. Nevertheless, determining how these components interact in living entities is a task that is still beyond the reach of the current laboratory methodologies. Understanding these interactions in the context of biological networks such as, e.g., cellular signaling pathways, is a relevant problem in biology.

Various approaches based on computer science have proven to be useful for addressing these issues. Formal models, for instance, allow us to make precise statements about the properties of biological systems, classifying them and, possibly, deducing other properties which are hard to discover by intuition or experimentally. It is worth noticing that, in general, the features of biochemical systems are often expressed informally, thus making it difficult or impossible to reason about them.

Several frameworks have been used for modeling various aspects of biological systems (see e.g., [1–5]). The characteristics of each model are mainly shaped by the features of the formalism it relies upon. Indeed, for a given model, some properties can be straightforwardly expressed while others must be abstracted away, due to the inherent limitations of the used language.

---

* Corresponding author.
E-mail addresses: carlos.olarte@gmail.com (C. Olarte), davide.chiarugi@gmail.com (D. Chiarugi), moreno.falaschi@unisi.it (M. Falaschi), dhermith@gmail.com (D. Hermith).

In this paper we focus on biological networks whose interactions depend on both time and spatial locations. We study how to give a declarative meaning to those interactions by providing encodings of them into intuitionistic linear logic (ILL) [6] with subexponentials [7–9] (SELL). ILL is a substructural logic where formulas are seen as resources and then, its proof system(s) controls the number of times a formula can be used during a proof. In SELL, such control is much richer and it allows us to define different modalities (e.g., locations) where formulas can be stored. Hence, the role of the subexponentials, as clarified later, is to specify the two biologically relevant main dimensions in our study, namely time and space. We show that our method is general enough to encode P systems [10,11], a general model of computation inspired by cells structures. We show that different semantics for P systems, such as minimal [12] and maximal parallelism [10] semantics, can be logically characterized in the same framework.

The encodings we propose are shown to have a strong level of adequacy. More precisely, biological steps correspond one-to-one to (focused [13]) derivations in the SELL system. Hence, a proof of a given property can be directly traced to steps in the biological system. Our proof-theoretic characterization of spatial and timed dependencies in biochemical systems thus has at least two main benefits:

- From the biological point of view, our encodings open the possibility to use all the meta theory and tools developed for linear logic to specify and verify biochemical systems featuring spatial and temporal modalities. One salient characteristic of our approach is that both, the system and its properties, are specified in the same logical framework. This is particularly relevant since in many works related to ours (see Sections 5 and 6 for details), usually one formalism is used for specifying the model while at least one different formalism is used for expressing the properties of interest (e.g., a temporal logic) and for proving them (e.g., by using a model checker).
- From the computational point of view, we give a further step in showing that linear logic (with subexponentials) is a general framework to specify and verify concurrent systems. Other studies relating concurrent formalisms and linear logic can be found, e.g., in [14–17,8].

A preliminary short version of this paper appeared in [18]. Here, we significantly review, enhance and refine our previous work. In particular, we present the full set of proofs, we clarify crucial technical details and we introduce many more explanations and examples, including an application of our framework in the context of the TWEAK-Fn14 cell signaling pathway [19].

The rest of the paper is structured as follows. In Section 2 we recall some concepts about subexponentials in linear logic. Section 3 defines an encoding of biochemical reactions that considers spatial and temporal modalities. We also show how to exploit the underlying logic for expressing some properties of the system. Section 4 provides a logical characterization of P systems as SELL formulas. Section 5 highlights how the proof theory of SELL may be used to draw conclusions about the studied biochemical systems. Section 6 discusses related work and concludes the paper.

2. Linear logic with subexponentials

In this section we recall the proof theory of intuitionistic linear logic (ILL) [6] with subexponentials [7,8]. Although this review should suffice to understand the developments in the forthcoming sections, we assume that the reader is familiar with logic and proof theory (see e.g., [20]).

Linear logic [6] is a substructural logic where formulas can be seen as resources. Hence, there is an explicit control over the number of times a formula can be used in a proof. More precisely, formulas can be split into two sets: classical (those that can be used as many times as needed) or linear (those that are consumed after being used). Classical formulas are marked with the modal operators !, ?, called exponentials. For instance, the formula $!F$ allows F to be used arbitrarily many times.

Intuitionistic linear logic with subexponentials [7] (SELL) shares with linear logic all its connectives except the exponentials. The subexponentials ($!^a$, $?^a$) add an index to the exponentials, thus allowing for splitting the formulas into many sets, each of which can then be specified to be classical (i.e., unbounded) or linear. In this case, the formula $!^a F$ can be interpreted as F holding in a given modality a, e.g., in the space location a.

As we shall see, the subexponentials provide a finer control on proofs and they allow for the specification of different modalities such as time-units or spatial locations. In this paper we shall use the system $SELL^\cap$ proposed in [8,9] that enjoys good proof theoretic properties: it admits cut-elimination and a sound and complete focused proof system [8,9]. Focusing [13] is a discipline on proofs to reduce the non-determinism during proof search. Hence, focused proofs can be interpreted as the normal form proofs for proof search. Although the proof rules of the focused system may look more involved at first glance, we shall rely on them to prove the adequacy results in Section 3. We thus introduce here the focused system and, for the sake of readability, we simplify a bit the notation. We also confine ourselves to the proof rules needed in the forthcoming sections, for instance, we do not introduce the rules for $?^a$, a connective not used in our encodings. The reader may refer to [8,9] for deeper technical details.

Connectives are separated into *negative* $\multimap, \&, \top, \forall, \cap, \top$ and *positive* $\otimes, \oplus, \exists, \cup, !^s$. The polarity of non-atomic formulas is inherited from its outermost connective (e.g., $F \multimap G$ is a negative formula while $F \otimes G$ is a positive one). Although the bias assigned to atoms does not interfere with provability [21], it changes considerably the shape of proofs. Here we require atoms to have a positive behavior.