Automatically deriving ODEs from process algebra models of signalling pathways

Muffy Calder[1], Stephen Gilmore[2], and Jane Hillston[2]

[1] Department of Computing Science, The University of Glasgow, Glasgow, Scotland
muffy@dcs.gla.ac.uk
[2] LFCS, School of Informatics, The University of Edinburgh, Edinburgh, Scotland
{stg,jeh}@inf.ed.ac.uk

Abstract. Differential equations are a classical approach for biochemical system modelling and have frequently been used to describe reactions of interest in biochemical pathways. Process algebras have also been applied in a small number of cases to describe such systems. In this paper we establish a connection between these approaches. This has the benefit of allowing process algebra models to be validated against trusted ODEs or, conversely, allowing ODEs derived from process algebra models to be evaluated and compared using bisimulation or other methods. In addition the process algebra models may now be efficiently solved using numerical differential equations procedures such as adaptive fifth-order Runge-Kutta.

1 Introduction

In recent years there has been some interest, and success, in applying formal system description techniques which originate in theoretical computer science to modelling biomolecular systems [9, 10, 7]. These description formalisms come equipped with apparatus to manipulate and reason about descriptions and formally extract underlying mathematical models. Thus analysis may be carried out in a rigorous manner. This is in contrast to mathematical models being developed directly, based on experimental data and modeller experience, but without formal underpinning.

In the case of one class of formalisms originating in computer science, *process algebras*, work so far has been focused on deriving stochastic simulations from the system description, based on Gillespie's algorithm [6]. In this paper we present an alternative use of process algebra models: to automatically generate systems of ordinary differential equations (ODEs) for models of intracellular signalling pathways. In some circumstances this is still the mathematical model of preference. There are several advantages to be gained by introducing a process algebra model as an intermediary to the derivation of the ODEs.

– The formal nature of the process algebra means that it is relatively straightforward to write a program to automatically generate the equivalent set of ODEs (one for each substrate), thus reducing the potential for human error.

Indeed we have already done so. Many signalling pathways have tens of substrates, resulting in tens of ODEs within the model. It can be challenging to accurately derive such a system by hand.
- Furthermore the formality of the process algebra model and its underlying semantics allow us to derive properties of the model, such as freedom from deadlock, before numerical analysis is carried out.
- Finally, an algebraic formulation of the model makes clear the interactions between the biochemical entities, or substrates. This is not always apparent in the classical ODE models. The style of modelling is descriptive, closely related to informal graphical representations that biochemists already use. Thus a change in the hypothesised network can often be much more readily made in the process algebra model than in the set of ODEs, where the change may have a pervasive impact.

The process algebra which we use is Hillston's PEPA [8], a Markovian process algebra which incorporates activity durations and probabilistic choices.

The most fundamental cellular processes are controlled by extracellular signalling [5]. This signalling, or communication between cells, is based upon the release of signalling molecules, which migrate to other cells and deliver stimuli to them (e.g. protein phosphorylation). Cell signalling is of special interest to cancer researchers because when cell signalling pathways operate abnormally, cells divide uncontrollably.

The remainder of the paper is structured as follows. In the following section we introduce the process algebra which we use, PEPA and demonstrate its use to describe a simple synthetic system. In Section 3 we explain how a system of ODEs may be derived from suitable PEPA models. We demonstrate the technique on a larger, realistic network in Section 4 and we conclude and offer some perspectives on future work in Section 5.

2 PEPA

Primarily, PEPA has been used to determine performance-related problems such as bottlenecks and hotspots in the design of information systems. As in all process algebras, systems are represented as the composition of *components* or *agents* which undertake *actions*. In PEPA the actions are assumed to have a duration, or delay. Thus the expression $(\alpha, r).P$ denotes a component which can undertake an α action, at rate r to evolve into a component P. PEPA is termed a *Markovian* process algebra because the duration associated with an action is usually assumed to be a random variable with a negative exponential distribution. Thus, r is the parameter of the corresponding distribution function ($F(t) = 1 - e^{-rt}$). However, as we will see in this paper, other interpretations of the rate information are also possible.

PEPA has a small set of combinators, allowing system descriptions to be built up as the concurrent performance and interaction of simple sequential components. We informally introduce the syntax below. More detail can be found in [8].

Prefix: The basic mechanism for describing the behaviour of a system with a PEPA model is to give a component a designated first action using the prefix combinator, denoted by a full stop, which was introduced above. As explained, $(\alpha, r).P$ carries out an α action with rate r, and it subsequently behaves as P.

Choice: The component $P + Q$ represents a system which may behave either as P or as Q. The activities of both P and Q are enabled. The first activity to complete distinguishes one of them: the other is discarded. The system will behave as the derivative resulting from the evolution of the chosen component.

Constant: It is convenient to be able to assign names to patterns of behaviour associated with components. Constants are components whose meaning is given by a defining equation. The notation for this is $X \stackrel{def}{=} E$. The name X is in scope in the expression on the right hand side meaning that, for example, $X \stackrel{def}{=} (\alpha, r).X$ performs α at rate r forever.

Hiding: The possibility to abstract away some aspects of a component's behaviour is provided by the hiding operator, denoted P/L. Here, the set L identifies those activities which are to be considered internal or private to the component and which will appear as the unknown type τ.

Cooperation: We write $P \underset{L}{\bowtie} Q$ to denote cooperation between P and Q over L. The set which is used as the subscript to the cooperation symbol, the *cooperation set* L, determines those activities on which the *cooperands* are forced to synchronise. For action types not in L, the components proceed independently and concurrently with their enabled activities. We write $P \parallel Q$ as an abbreviation for $P \underset{L}{\bowtie} Q$ when L is empty.

However, if a component enables an activity whose action type is in the cooperation set it will not be able to proceed with that activity until the other component also enables an activity of that type. The two components then proceed together to complete the *shared activity*. The rate of the shared activity may be altered to reflect the work carried out by both components to complete the activity.

In some cases, when an activity is known to be carried out in cooperation with another component, a component may be *passive* with respect to that activity. This means that the rate of the activity is left unspecified (denoted $\top$) and is determined upon cooperation, by the rate of the activity in the other component. All passive actions must be synchronised in the final model.

2.1 Using PEPA to model intracellular signalling pathways

In [1] we investigated the use of PEPA and Markov process analysis to study the ERK signalling pathway. In particular we considered the issue of how to represent concentrations and presented two distinct styles of PEPA model. In the first, each component of the PEPA model corresponds to a substrate of the pathway. The possible range of concentrations is discretized and representative rates chosen to span a subrange of concentrations. In [1] we take the coarsest possible discretization considering only *high* and *low* concentrations, in which low concentrations are assumed to be unable to participate in any reactions. The second

style of model focuses on sub-processes within the pathway. For each substrate known to have an initially high concentration, one PEPA component represents its possible evolution through a number of reactions and compounds. When intermediate levels of concentration are required, multiple instances of pathway components may be used. In [2] these two styles of model are shown to give rise to equivalent underlying representations, differing only in the description style. Moreover systematic transformations between them are specified. In this paper we only consider the substrate models, knowing that the other representation is equivalent and can be derived.

To illustrate our ideas we consider a small synthetic example pathway shown in Figure 1. This is chosen because of its compact size, facilitating an accessible comparison between the process algebra and ODE views of the pathway.

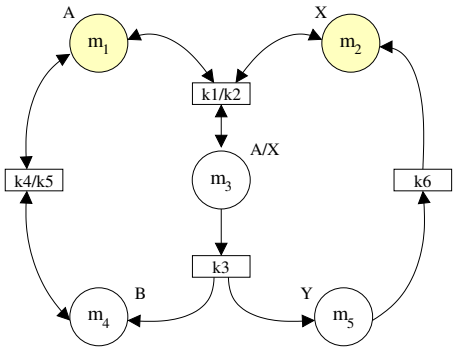


Fig. 1. Small synthetic example network

In this network we assume that there are five reactants (two sub-pathways) stemming from initial concentrations in substrates A and X. The diagram can be understood as follows: substrates A and X can associate with rate constant $k1$ to form compound A/X which disassociates into A and X with rate constant $k2$ or forms the products B and Y with rate constant $k3$. B and Y become A and X respectively with rates $k4$ and $k6$ respectively, the $B \longrightarrow A$ reaction being reversible at rate $k5$.

The reactant-based PEPA model has the following form, where the subscripts "H" and "L" denote high and low concentrations respectively:

$$A_H \stackrel{def}{=} (k1react, k1).A_L + (k5react, k5).A_L$$
$$A_L \stackrel{def}{=} (k2react, k2).A_H + (k4react, k4).A_H$$

$$X_H \stackrel{def}{=} (k1react, k1).X_L$$
$$X_L \stackrel{def}{=} (k2react, k2).X_H + (k6react, k6).X_H$$

$$A/X_H \stackrel{def}{=} (k2react, k2).A/X_L + (k3react, k3).A/X_L$$
$$A/X_L \stackrel{def}{=} (k1react, k1).A/X_H$$

$$B_H \stackrel{def}{=} (k4react, k4).B_L$$
$$B_L \stackrel{def}{=} (k5react, k5).B_H + (k3react, k3).B_H$$

$$Y_H \stackrel{def}{=} (k6react, k6).Y_L$$
$$Y_L \stackrel{def}{=} (k3react, k3).Y_H$$

The complete model of the network is the interation of these components constrained by cooperation to share the appropriate actions:

$$(((A_{H} \underset{\{k1react,k2react\}}{\bowtie} X_H) \underset{\{k1react,k2react\}}{\bowtie} A/X_L) \underset{\{k3react,k4react,k5react\}}{\bowtie} B_L) \underset{\{k3react,k6react\}}{\bowtie} Y_L$$

3 Automatically deriving ODEs

Even at the coarsest level of abstraction, distinguishing only high and low concentrations the reactant-based model provides sufficient information for deriving an ODE representation of the same system. It is sufficent to know which reactions increase concentration (low-to-high) and which decrease it (high-to-low).

For any reactant-based PEPA model with derivatives designated high and low, it is straightforward to construct an *activity graph* which captures this information.

Definition 1 (Activity Graph). *An* activity graph *is a bipartite graph* (N, A). *The nodes* N *are partitioned into* N_r*, the* reactions*, and* N_a*, the* reagents. $A \subset (N_r \times N_a) \cup (N_a \times N_r)$*, where* $a = (n_r, n_a) \in A$ *if* n_r *is a reaction in which the concentration of reagent* n_a *is increased, and* $a = (n_a, n_r) \in A$ *if* n_r *is a reaction in which the concentration of reagent* n_a *is decreased.*

The same information can be represented in a matrix, termed the *activity matrix*.

Definition 2 (Activity Matrix). *For a pathway with* R *reactions and* S *reagents, the* activity matrix M_a *is an* $S \times R$ *matrix, and the entries are defined as follows.*

$$(s_i, r_j) = \begin{cases} +1 & \text{if } (r_j, s_i) \in A \\ -1 & \text{if } (s_i, r_j) \in A \\ 0 & \text{if } (s_i, r_j) \notin A \cup (r_j, s_i) \notin A \end{cases}$$

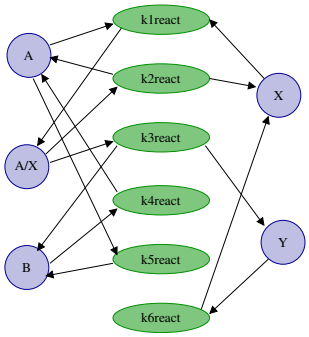
In the activity matrix each row corresponds to a single reactant[3]. In the representation of the pathway as a systems of ODEs there is one equation for each reactant, detailing the impact of the rest of the system on the concentration of that reactant. This can derived automatically from the activity matrix, when

[3] The activity matrix is clearly related to the stochiometry matrix.

we associate a concentration variable m_i with each row of the matrix. The entries in the row indicate which reactions have an impact on this reactant, the sign of the entry showing whether the effect is to increase or decrease concentration. Thus the number of terms in the ODE will be equal to the number of non-zero entries in the corresponding row, each term being based on the rate constant for the reaction associated with that row. By the law of mass action, the actual rate of change caused by each reaction will be the rate constant multiplied by the current concentration of those reactants consumed in the reaction. The identity of these reactants can be found in the column corresponding to the reaction, a negative entry indicating that a reactant is consumed.

3.1 Small example revisited

The activity graph and activity matrix corresponding to the reactant-based PEPA model of the small example network shown in Figure 1 are shown in Figure 2.



reactants	k1react	k2react	k3react	k4react	k5react	k6react	*concentration variables*
A	-1	$+1$	0	$+1$	-1	0	m_1
X	-1	$+1$	0	0	0	$+1$	m_2
A/X	$+1$	-1	-1	0	0	0	m_3
B	0	0	$+1$	-1	$+1$	0	m_4
Y	0	0	$+1$	0	0	-1	m_5

Fig. 2. Activity graph and activity matrix for the small example

Based on the matrix (in Figure 2) it is straightforward to derive the differential equations which are easily validated against the original system.

$$\frac{dm_1(t)}{dt} = -k1m_1(t)m_2(t) + k2m_3(t) + k4m_4(t) - k5m_1(t)$$
$$\frac{dm_2(t)}{dt} = -k1m_1(t)m_2(t) + k2m_3(t) + k6m_5(t)$$
$$\frac{dm_3(t)}{dt} = k1m_1(t)m_2(t) - k2m_3(t) - k3m_3(t)$$
$$\frac{dm_4(t)}{dt} = k3m_3(t) - k4m_4(t) + k5m_1(t)$$
$$\frac{dm_5(t)}{dt} = k3m_3(t) - k6m_5(t)$$

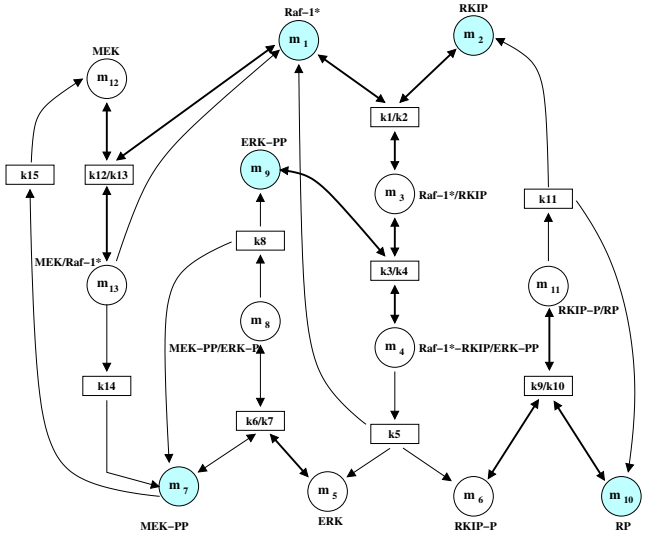


Fig. 3. RKIP inhibited ERK pathway

4 Case study: the ERK intracellular signalling pathway

The Ras/Raf-1/MEK/ERK pathway (also called Ras/Raf, or ERK pathway) is a ubiquitous pathway that conveys mitogenic and differentiation signals from the cell membrane to the nucleus. Briefly, Ras is activated by an external stimulus, it then binds to and activates Raf-1 (to become Raf-1*, "activated" Raf) which in turn activates MEK and then ERK. This "cascade" of protein interaction controls cell differentiation, the effect being dependent upon the activity of ERK. A current area of experimental scientific investigation is the role the kinase inhibitor protein RKIP plays in the behaviour of this pathway: the hypothesis is that it inhibits activation of Raf and thus can "dampen" down the ERK pathway. Certainly there is much evidence that RKIP inhibits the malignant transformation by Ras and Raf oncogenes in cell cultures and it is reduced in tumours. Thus good models of these pathways are required to understand the role of RKIP and develop new therapies. Moreover, an understanding of the functioning and structure of this pathway may lead to more general results applicable to other pathways.

Here, we consider the RKIP inhibited ERK pathway as presented in [4], based on the graphical representation given in Figure 3 (taken from [4], with some additions[4]).

We take Figure 3 as our starting point, and explain informally, its meaning. Each node is labelled by the protein (or substrate, we use the two interchangably)

[4] Analysis of our original model(s) indicated a problem with MEK and prompted us to contact an author of [4] who confirmed that there was an omission.

it denotes. For example, Raf-1, RKIP and Raf-1*/RKIP are proteins, the last being a complex built up from the first two. A sufffix -P or -PP denotes a phosyphorylated protein, for example MEK-PP and ERK-PP. Each protein has an associated concentration, denoted by $m1$, $m2$ etc. In the figure, bi-directional arrows denote both forward and backward reactions; uni-directional arrows denote disassociations. For example, Raf-1* and RKIP react (forwards) to form Raf-1*/RKIP, and Raf-1/RKIP disassociates (a backward reaction) into Raf-1* and RKIP. Each reaction has a rate denoted by the rate constants $k1$, $k2$, etc. These are given in the rectangles, with $kn/kn+1$ denoting that kn is the forward rate and $kn+1$ the backward rate. So for example, Raf-1* and RKIP react (forwards) with rate $k1$, and Raf-1/RKIP disassociates with rate $k2$. Initially, all concentrations are unobservable, except for m_1, m_2, m_7, m_9, and m_{10} [4].

4.1 Modelling the ERK signalling pathway in PEPA

The model we present is a reagent-centric view, focussing on the variations in concentrations of the reagents, fluctuating with phosphorylation and product formation, i.e. with association and disassociation reactions. This model provides a fine-grained, distributed view of the system. Each reaction in the pathway is represented by a multi-way synchronisation – on the reagents of the reaction[5].

$$
\begin{aligned}
\text{Raf-1}^*_{\text{H}} &\stackrel{def}{=} (k1react, k_1).\text{Raf-1}^*_{\text{L}} + (k12react, k_{12}).\text{Raf-1}^*_{\text{L}} \\
\text{Raf-1}^*_{\text{L}} &\stackrel{def}{=} (k5product, k_5).\text{Raf-1}^*_{\text{H}} + (k2react, k_2).\text{Raf-1}^*_{\text{H}} \\
&\quad + (k13react, k_{13}).\text{Raf-1}^*_{\text{H}} + (k14product, k_14).\text{Raf-1}^*_{\text{H}} \\
\\
\text{RKIP}_{\text{H}} &\stackrel{def}{=} (k1react, k_1).\text{RKIP}_{\text{L}} \\
\text{RKIP}_{\text{L}} &\stackrel{def}{=} (k11product, k_{11}).\text{RKIP}_{\text{H}} + (k2react, k_2).\text{RKIP}_{\text{H}} \\
\\
\text{ERK-PP}_{\text{H}} &\stackrel{def}{=} (k3react, k_3).\text{ERK-PP}_{\text{L}} \\
\text{ERK-PP}_{\text{L}} &\stackrel{def}{=} (k8product, k_8).\text{ERK-PP}_{\text{H}} + (k4react, k_4).\text{ERK-PP}_{\text{H}} \\
\\
\text{Raf-1}^*/\text{RKIP}_{\text{H}} &\stackrel{def}{=} (k3react, k_3).\text{Raf-1}^*/\text{RKIP}_{\text{L}} + (k2react, k_2).\text{Raf-1}^*/\text{RKIP}_{\text{L}} \\
\text{Raf-1}^*/\text{RKIP}_{\text{L}} &\stackrel{def}{=} (k1react, k_1).\text{Raf-1}^*/\text{RKIP}_{\text{H}} + (k4react, k_4).\text{Raf-1}^*/\text{RKIP}_{\text{H}} \\
\\
\text{Raf-1}^*/\text{RKIP/ERK-PP}_{\text{H}} &\stackrel{def}{=} (k5product, k_5).\text{Raf-1}^*/\text{RKIP/ERK-PP}_{\text{L}} \\
&\quad + (k4react, k_4).\text{Raf-1}^*/\text{RKIP/ERK-PP}_{\text{L}} \\
\text{Raf-1}^*/\text{RKIP/ERK-PP}_{\text{L}} &\stackrel{def}{=} (k3react, k_3).\text{Raf-1}^*/\text{RKIP/ERK-PP}_{\text{H}} \\
\vdots \qquad &\quad \vdots \qquad\qquad \vdots
\end{aligned}
$$

Fig. 4. PEPA model definitions for the reagent-centric model

[5] We agree with the authors of [9] – reactions are fundamentally synchronous.

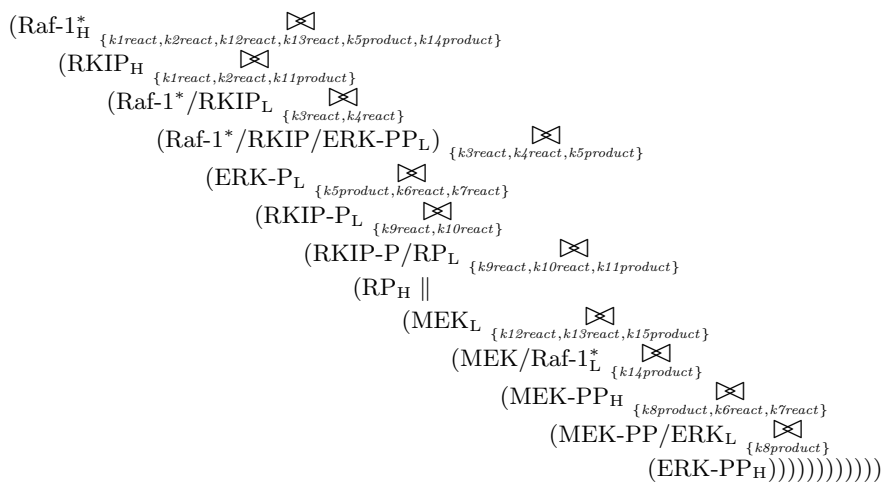


Fig. 5. PEPA model configuration for the reagent-centric model

The model is presented in Figures 4 and 5. For brevity we present only some of the reactant definitions. The interested reader is referred to [1] for the full model. We distinguish between high (i.e. observable) and low (i.e. unobservable) concentrations of reagents. The former implies that a reagent *can* participate (as a producer) in a forward reaction; the latter implies that a reagent *can* participate (as a consumer) in a product, or (as a producer) in a backward reaction. Otherwise, the substrate is inert, with respect to a reaction. We define the behaviour of each substrate in turn, for each concentration. Thus there are $2n$ equations, where n is the number of proteins. We maintain the naming convention that high concentrations have a H subscript and low concentrations have a L subscript.

The PEPA equation in Figure 5 shows how these reactant components are composed in order to obtain the permissible interleavings of reactions. It also defines the initial state of the model. This has high concentrations of some reagents and low concentrations of others, based on the experimental observations [4].

4.2 Deriving ODEs

From the PEPA model we can derive the activity matrix shown in Figure 6. The component definitions (Figure 4) determine which reactants are involved in which reactions, and the nature of the involvement (i.e. producer *vs.* consumer) while the system equation (Figure 5) establishes the synchronisations. In fact in this model all actvities of the same name are carried out in cooperation but this need not be the case in general.

As previously stated, each row corresponds to a single *reagent*; the entries in a row indicate whether an activity (column) increases the concentration (+1),

	$k1$	$k2$	$k3$	$k4$	$k5$	$k6$	$k7$	$k8$	$k9$	$k10$	$k11$	$k12$	$k13$	$k14$	$k15$
Raf-1*	−1	+1	0	0	+1	0	0	0	0	0	0	−1	+1	+1	0
RKIP	−1	+1	0	0	0	0	0	0	0	0	+1	0	0	0	0
Raf-1*/RKIP	+1	−1	−1	+1	0	0	0	0	0	0	0	0	0	0	0
Raf-1*/RKIP/ERK-PP	0	0	+1	−1	−1	0	0	0	0	0	0	0	0	0	0
ERK-P	0	0	0	0	+1	−1	+1	0	0	0	0	0	0	0	0
RKIP-P	0	0	0	0	+1	0	0	0	−1	+1	0	0	0	0	0
MEK-PP	0	0	0	0	0	−1	+1	+1	0	0	0	0	0	+1	−1
MEK-PP/ERK	0	0	0	0	0	+1	−1	−1	0	0	0	0	0	0	0
ERK-PP	0	0	−1	+1	0	0	0	+1	0	0	0	0	0	0	0
RP	0	0	0	0	0	0	0	0	−1	+1	+1	0	0	0	0
RKIP-P/RP	0	0	0	0	0	0	0	0	+1	−1	−1	0	0	0	0
MEK	0	0	0	0	0	0	0	0	0	0	0	−1	+1	0	+1
MEK/Raf-1*	0	0	0	0	0	0	0	0	0	0	0	+1	−1	−1	0

Fig. 6. Activity matrix of the ERK pathway

decreases it (-1) or has no impact (0). Each column corresponds to a single *reaction*; the negative entries indicate those substrates which are *producers* (and therefore consumed) in the reaction. One ODE is then derived from each row of the matrix.

For example, if we consider the compound Raf-1*/RKIP, in the third row of the matrix. It is involved in four reactions (indicated by the bold entries).

	k1	**k2**	**k3**	**k4**	$k5$	$k6$	...	*conc.*
Raf-1*	*-1*	+1	0	0	+1	0	...	m_1
RKIP	*-1*	+1	0	0	0	0	...	m_2
Raf-1*/RKIP	**+1**	***-1***	***-1***	**+1**	0	0	...	$\mathbf{m_3}$
Raf-1*/RKIP/ERK-PP	0	0	+1	*-1*	−1	0	...	m_4
ERK-P	0	0	0	0	+1	−1	...	m_5
RKIP-P	0	0	0	0	+1	0	...	m_6
MEK-PP	0	0	0	0	0	−1	...	m_7
MEK-PP/ERK	0	0	0	0	0	+1	...	m_8
ERK-PP	0	0	*-1*	+1	0	0	...	m_9
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱	

For each of those reactions the producers are indicated by the italic entries in the corresponding column, resulting in the following differential equation for the concentration of Raf-1*/RKIP:

$$\frac{dm_3(t)}{dt} = k1\, m_1(t)m_2(t) - k2\, m_3(t) - k3\, m_3(t)m_9(t) + k4\, m_4(t)$$

The remaining set of automatically derived equations are shown below.

$$
\begin{aligned}
\frac{dm_1(t)}{dt} &= -k_1 m_1(t) m_2(t) + k_2 m_3(t) + k_5 m_4(t) - k_{12} m_1(t) m_{12}(t) \\
&\quad + k_{13} m_{13}(t) + k_{14} m_{13}(t) \\
\frac{dm_2(t)}{dt} &= -k_1 m_1(t) m_2(t) + k_2 m_3(t) + k_{11} m_{11}(t) \\
\frac{dm_4(t)}{dt} &= k_3 m_3(t) m_9(t) - k_4 m_4(t) - k_5 m_4(t) \\
\frac{dm_5(t)}{dt} &= k_5 m_4(t) - k_6 m_5(t) m_7(t) + k_7 m_8(t) \\
\frac{dm_6(t)}{dt} &= k_5 m_4(t) - k_9 m_6(t) m_{10}(t) + k_{10} m_{11}(t) \\
\frac{dm_7(t)}{dt} &= -k_6 m_5(t) m_7(t) + k_7 m_8(t) + k_8 m_8(t) + k_{14} m_{13}(t) - k_{15} m_7(t) \\
\frac{dm_8(t)}{dt} &= k_6 m_5(t) m_7(t) - k_7 m_8(t) - k_8 m_8(t) \\
\frac{dm_9(t)}{dt} &= -k_3 m_3(t) m_9(t) + k_4 m_4(t) + k_8 m_8(t) \\
\frac{dm_{10}(t)}{dt} &= -k_9 m_6(t) m_{10}(t) + k_{10} m_{11}(t) + k_{11} m_{11}(t) \\
\frac{dm_{11}(t)}{dt} &= k_9 m_6(t) m_{10}(t) - k_{10} m_{11}(t) - k_{11} m_{11}(t) \\
\frac{dm_{12}(t)}{dt} &= -k_{12} m_1(t) m_{12}(t) + k_{13} m_{13}(t) + k_{15} m_7(t) \\
\frac{dm_{13}(t)}{dt} &= k_{12} m_1(t) m_{12}(t) - k_{13} m_{13}(t) - k_{14} m_{13}(t)
\end{aligned}
$$

5 Conclusions

Stochastic process algebras have found new applications in modelling biochemical pathways. In addition to quantified analysis, such models offer facilities to reason about the system model and investigate its structural properties [3]. Previously quantified analysis was carried out via simulation or procedures of numerical linear algebra, both of which are computationally expensive processes and do not scale well to allow the representation of realistic populations of reactants. Analysis based on the use of differential equations has until now been unavailable to stochastic process algebra models. The contribution of this paper is to establish a bridge between the two approaches.

Representing the system in process algebra has several tangible benefits. The compositional nature of the system description makes it easy to make a change in the hypothesised role of a reagent within a network. In general this will involve changing only the expressions representing the behaviour of this reagent, whereas the impact on the ODEs may be pervasive. As shown in [2] the reagent- and pathway-centric PEPA models have complementary strengths. In

particular the pathway models capture structural information which is lost in the ODE representation.

The method presented here for the derivation of ODEs from the process algebra model is fully automatic and has been implemented in order that we may test the effectiveness of the method on models of larger scale. We have used a fifth-order Runge-Kutta solver together with a tool for the PEPA language to analyse the PEPA ERK pathway model presented in this paper.

Acknowledgements

Muffy Calder is supported by the DTI Beacon Bioscience Projects programme. Stephen Gilmore and Jane Hillston are supported by the DEGAS (Design Environments for Global ApplicationS) project IST-2001-32072.

References

1. M. Calder, S. Gilmore, and J. Hillston. Modelling the influence of RKIP on the ERK signaling pathway using the stochastic process algebra PEPA. In *Proceedings of BioConcur'04*, London, England, August 2004.
2. M. Calder, S. Gilmore, and J. Hillston. Modelling the influence of RKIP on the ERK signaling pathway using the stochastic process algebra PEPA. Extended version of BioConcur'04 paper, submitted for publication, 2005.
3. M. Calder, V. Vyshemirsy, D. Gilber, and R. Orton. Analysis of signalling pathways using prism model checker. In *Proc. of Computational Methods in Systems Biology Workshop*, 2005.
4. K.-H. Cho, S.-Y. Shin, H.-W. Kim, O. Wolkenhauer, B. McFerran, and W. Kolch. Mathematical modeling of the influence of RKIP on the ERK signaling pathway. In C. Priami, editor, *Computational Methods in Systems Biology (CSMB'03)*, volume 2602 of *LNCS*, pages 127–141. Springer-Verlag, 2003.
5. W.H. Elliot and D.C. Elliot. *Biochemistry and Molecular Biology, 2nd edition.* Oxford University Press, 2002.
6. D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
7. M. Heiner and I. Koch. Petri net based model validation in systems biology. In *25th International Conference on Application and Theory of Petri Nets, Bologna, Italy*, 2004.
8. J. Hillston. *A Compositional Approach to Performance Modelling.* Cambridge University Press, 1996.
9. C. Priami, A. Regev, W. Silverman, and E. Shapiro. Application of a stochastic name passing calculus to representation and simulation of molecular processes. *Information Processing Letters*, 80:25–31, 2001.
10. A. Regev, W. Silverman, and E. Shapiro. Representation and simulation of biochemical processes using the pi-calculus process algebra. In *Proceedings of the Pacific Symposium of Biocomputing (PSB2001)*, pages 459–470, 2001.