Modelling and Analysis of Biochemical Signalling Pathway Cross-talk

Robin Donaldson, Muffy Calder

Department of Computing Science,
University of Glasgow
Glasgow, UK

radonald@dcs.gla.ac.uk,

muffy@dcs.gla.ac.uk

Signalling pathways are abstractions that help life scientists structure the coordination of cellular activity. Cross-talk between pathways accounts for many of the complex behaviours exhibited by signalling pathways and is often critical in producing the correct signal-response relationship. Formal models of signalling pathways and cross-talk in particular can aid understanding and drive experimentation. We define an approach to modelling based on the concept that a pathway is the (synchronising) parallel composition of instances of generic modules (with internal and external labels). Pathways are then composed by (synchronising) parallel composition and renaming; different types of cross-talk result from different combinations of synchronisation and renaming. We define a number of generic modules in PRISM and five types of cross-talk: signal flow, substrate availability, receptor function, gene expression and intracellular communication. We show that Continuous Stochastic Logic properties can both detect and distinguish the types of cross-talk. The approach is illustrated with small examples and an analysis of the cross-talk between the TGF-β/BMP, WNT and MAPK pathways.

1 Introduction

Signalling pathways[1] are well-known abstractions that help life scientists structure the coordination of cellular activity. Interaction between pathways, known as cross-talk, appears to have arisen for several reasons, for example to integrate signals, to produce a variety of responses to a signal, to reuse proteins between pathways; it accounts for many of the complex signalling behaviours.

This paper investigates modelling and analysis of pathway cross-talk; a key outcome is the suitability of process algebraic operators to modelling cross-talk. We define an approach to modelling based on the concept that a pathway is the (synchronising) parallel composition of instances of generic modules (with internal and external labels). Pathways are then composed by (synchronising) parallel composition and renaming; different types of cross-talk result from different combinations of synchronisation and renaming. We use the PRISM modelling language and model checker. The contribution of the paper is the following:

- novel categorisation of types of cross-talk

- modelling based on pathway module instantiation, internal/external reactions & synchronisation over subsets of (possibly renamed) external reactions

---

[1] we refer simply to pathways henceforth

- examples of modelling each type of cross-talk

- detection and characterisation of cross-talk using Continuous Stochastic Logic

- application of approach to a case study: cross-talk between TGF-β/BMP, WNT and MAPK pathways.

This paper is organised as follows. The following section outlines the background to signalling pathway cross-talk and defines a novel categorisation based on examples in the literature. In section 3 we describe the modelling approach. In section 4 we show how each type of cross-talk is modelled for an example pathway and how we can detect and characterise the cross-talk using CSL. Section 5 describes the case study, cross-talk between TGF-β/BMP, WNT and MAPK pathways, and gives some results. Section 6 contains a discussion of our overall approach and Section 7 reviews related work. Conclusions and directions for future work are in section 8.

Throughout the paper the following notation is used. Transformation (e.g. protein X turns into protein Y) is denoted by a solid line with an arrow. Catalysis (increase in the rate of a reaction) is denoted by a dashed line with an arrow. Inhibition (decrease in the rate of a reaction) is denoted by a solid line with a blunt end. Finally, we distinguish between inactive proteins and active proteins rather than the various mechanisms by which a protein changes state. An active protein is decorated with $*$, e.g X is inactive and $X*$ is active. This notation is illustrated in Figure 1(a).
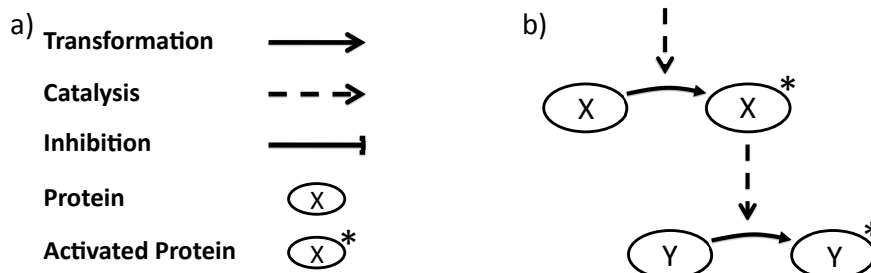

Figure 1: a) the notation used throughout this paper for arcs and nodes, b) an example of a 2-stage signalling cascade in which the activated protein X catalyses the activation of protein Y.

2   Signalling Pathways & Cross-talk

Pathways are the mechanism by which a cell receives a signal and produces the appropriate cellular response. A cell can respond to many different signals, and the same signalling pathway can produce different responses based on different signals. For example, the well-studied MAPK/ERK pathway can respond with cellular proliferation or cellular differentiation depending on the type of growth-factor present [1]. Many of the reactions involved in signalling pathways are enzyme catalysed protein activations, often arranged in a "signalling cascade". In such a cascade, the activated protein on one "level" is the enzyme for the activating reaction of the next "level", as shown in Figure 1(b).

Signalling pathways were first thought to be a linear series of reactions, but more recently, detailed understanding of these pathways shows that they are non-linear [2]. The series of reactions forming

the pathway can, for example, diverge or interact upstream/downstream in the chain of reactions forming a feedback/feedforward loop. Laboratory techniques measuring the concentration of proteins or RNAs/DNAs in the cell under different conditions now enable biologists to build appropriate abstractions of pathways. This is an incremental process further complicated by the lack of clear definition of what constitutes a single pathway. It is an inexact science to define the boundaries between pathways. Often, the boundaries are simply drawn such that the abstraction of the pathway can explain the biological data.

The term cross-talk was first applied to electronic circuits to describe a signal in one circuit having an undesired effect on another circuit [3]. Cross-talk in this setting is a design flaw: the electronic circuit has been specified and built, and has resulted in an undesired interaction between signals, called "signal interference". Biochemical cross-talk [4] is an interaction between signals flowing through two or more signalling pathways in a cell, however, this is not necessarily indicative of signal interference.

2.1 Types of Cross-talk

Cross-talk can occur at all stages of signal propagation through a pathway. Although there is some discussion of types of cross-talk [5], there appears to be no universal categorisation in the literature. Here, we define five categories of cross-talk: signal flow, substrate availability, receptor function, gene expression and intracellular communication. The five categories are illustrated in Figure 2 and are discussed in more detail below, with reference to an indicative example. We note that four of the five categories are alluded to in [6] but are not made specific.

2.1.1 Signal Flow Cross-talk

Signal flow cross-talk between two pathway occurs when a molecular species in one pathway affects the signal flow (rate of protein activation) in another pathway, shown in Figure 2(a). The species affects the signal flow by altering the rate(s) of the activation or deactivation reactions in the other pathway. Typically, this interaction occurs in the cytoplasm and affects the rate of the downstream signal. For example, in [4] there is signal flow cross-talk between the MAPK and Integrin signalling pathways. Activation of the Integrin pathway enhances signalling through the MAPK pathway by increased rate of activation of key proteins in the pathway.

2.1.2 Substrate Availability Cross-talk

Substrate availability cross-talk occurs when two pathways compete for one or more common or homologous proteins (proteins that perform the same function), shown in Figure 2(b). For example [7] describes two pathways that compete for activation of the MAPK cascade. The pathways share the MAPKKK protein STE11 and have homologous MAPKK and MAPK proteins.

2.1.3 Receptor Function Cross-talk

Cross-talk can occur at the level of the pathway receptor, shown in Figure 2(c). The receptor's ability to detect a ligand can be affected by cross-talk, for example, being inhibited (thus slowing or blocking signal propagation) or activated (thus producing signal propagating in absence of the ligand). In [8] other signalling pathways can activate the estrogen receptor in the absence of the estrogen ligand.

2.1.4   Gene Expression Cross-talk

Cross-talk can occur at the level of gene expression within the nucleus, shown in Figure 2(d). In this case, there is an interaction between pathways in terms of genes that will be expressed or repressed, for example, through activating/deactivating transcription factors (TFs) or increasing/reducing the number of TFs. In [9] two pathways contain cross-talk within the nucleus. One pathway contains a transcription factor called GR that resides outside the nucleus in its inactive state. Upon activation (signalling), GR relocates to the nucleus and represses the transcription factor NF-κB that is activated through another pathway.

2.1.5   Intracellular Communication Cross-talk

Finally, signalling pathways can cross-talk using the less direct manner of intracellular communication, shown in Figure 2(e). Instead of physical interaction between the proteins that comprise the pathways, one pathway can release a ligand that activates another pathway. In [5] the TGF-β/BMP and WNT pathways reciprocally regulate the production of their ligands. There is some contention in the literature as to whether this is genuine cross-talk: the interaction is less-direct than other types of cross-talk and involves lengthy processes such as gene expression and ligand excretion.

3   Modelling

Our aim is to build models of pathways and their cross-talk in a modular fashion, whilst demonstrating our cross-talk categorisation.

The literature contains many examples of modelling biological systems based on an unstructured set of equations, for example Ordinary Differential Equations (ODEs) [10]. However, flat equations have disadvantages such as lack of structure and modularity – there is no accepted way of dealing the with composition of these models. Several formal languages from Computer Science have well-understood notions of structure.

We choose a state-based formalism, the PRISM modelling language [11], with modules, renaming and synchronous communication between modules. The semantics is given by continuous-time Markov chains (CTMCs), which can be analysed using the PRISM model checker with Continuous Stochastic Logic (CSL). We give a brief overview of the language.

The language is based on "reactive modules", each of which can contain local variables. The variables are updated by the execution of commands that have the following syntax, `[label] guard -> rate:update_statement`. Commands are only executable when their `guard` becomes true, and any `label` synchronises as required. The `rate` is used to build the underlying CTMC, providing both the probability and timing information of the state after `update_statement` executes.

Modules can be composed concurrently, synchronising (multiway) on the commands whose labels occur in the synchronisation set, *L*. For example, given modules $M1$ and $M2$, and set of labels *L*, $M1 \,|[L]|\, M2$ denotes the concurrent composition of $M1$ and $M2$, synchronising on all labels in L. If the label set is omitted, $M1 \,||\, M2$, then $M1$ synchronises with $M2$ on the intersection of labels occurring in $M1$ and $M2$. PRISM also allows renaming of labels, denoted thus $M1\,\{old\_label \leftarrow new\_label\}$, and hiding, denoted thus $M\,/\,\{label1,\,\ldots,\,labeln\}$. Hidden labels are not available for synchronisation.
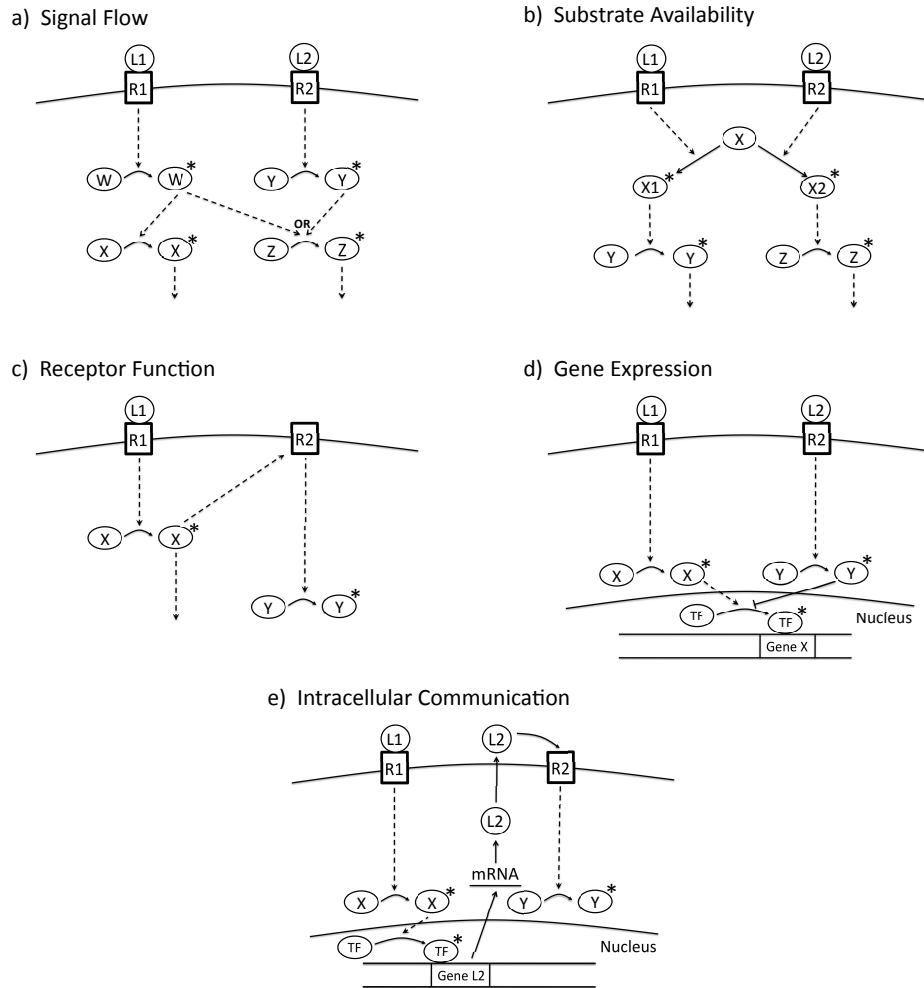
Figure 2: An example of each of the five types of cross-talk: a) a pathway up-regulates signal flow through another pathway, b) two pathways compete for a protein, c) a pathway activates the receptor of another pathway in the absence of a ligand, d) two pathways have conflicting transcriptional responses, e) a pathway releases a ligand for another pathway. The following notation is used: L1/L2 - ligands, R1/R2 - receptors, W/X/Y/Z - Proteins, TF - Transcription Factor.

3.1 Generic Modules

We define a *pathway module* to be a behavioural pattern within a pathway. For example, commonly occurring pathway modules are: receptor, 3-stage cascade and gene expression (Figure 3).

We represent these by generic modules in PRISM as follows. We adopt the reagent-centric modelling style, as first presented in [12]. Local variables contain the concentration of proteins in a particular state (e.g. active, inactive). Labelled commands change protein concentration according to biochemical reactions. Reactions are considered to be external or internal. The former denote behaviour that can coordinate (or be coordinated with) behaviour in other modules, they are available for synchronisation. The latter reactions are hidden and are not available for synchronisation. Concentration is modelled by
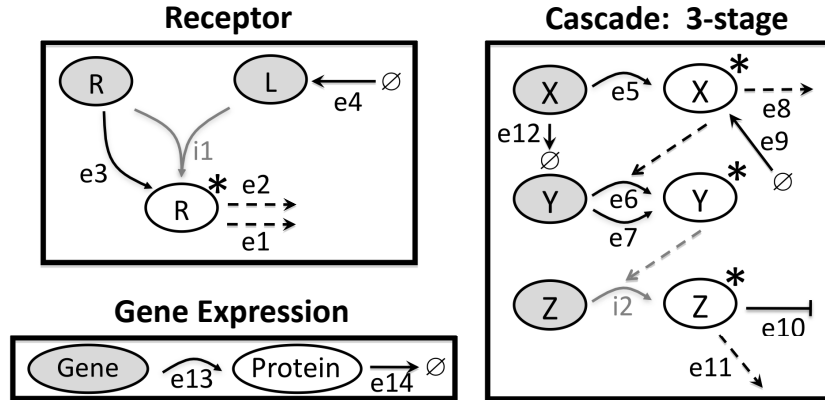
Figure 3: Three generic pathway modules: receptor, 3-stage cascade and gene expression. There are 14 external reactions ($e1,\ldots,e14$) denoted by black arcs, 2 internal reactions ($i1,i2$) denoted by gray arcs and 6 species that are present in the initial state denoted by shaded ellipses.

discrete, abstract *levels*, as defined in [13]. Here, we consider strong abstractions, usually two levels $\{0,1\}$, and unit reaction rates. Due to the ligand production reaction $e4$, ligand has three levels $\{0,1,2\}$, and the rate of ligand-receptor binding reflects the level of the ligand (law of mass-action).

The PRISM modules for receptor, 3-stage cascade and gene expression are the following.

```
module Receptor
   R : [0..1] init 1;    L1 : [0..2] init 1;    R1Active : [0..1] init 0;

   [i1_1] R1 = 1 & L1 >= 1 & R1Active = 0 -> L1:(R1' = 0) & (L1' = 0) & (R1Active' = 1);
   [e1_1] R1Active = 1 -> 1:(R1Active' = R1Active);
   [e2_1] R1Active = 1 -> 1:(R1Active' = R1Active);
   [e3_1] R1 = 1 & R1Active = 0 -> 1:(R1' = 0) & (R1Active' = 1);
   [e4_1] L1 < 2 -> 1:(L1' = L1 + 1);
endmodule

module Cascade3
   X1Inactive : [0..1] init 1;    X1Active : [0..1] init 0;    Y1Inactive : [0..1] init 1;
   Y1Active : [0..1] init 0;    Z1Inactive : [0..1] init 1;  Z1Active : [0..1] init 0;

   [e5_1] X1Inactive = 1 & X1Active = 0 -> 1:(X1Inactive' = 0) & (X1Active' = 1);
   [e6_1] Y1Inactive = 1 & Y1Active = 0 & X1Active = 1 -> 1:(Y1Inactive' = 0) & (Y1Active' = 1);
   [e7_1] Y1Inactive = 1 & Y1Active = 0 -> 1:(Y1Inactive' = 0) & (Y1Active' = 1);
   [i2_1] Z1Inactive = 1 & Z1Active = 0 & Y1Active = 1 -> 1:(Z1Inactive' = 0) & (Z1Active' = 1);
   [e8_1] X1Active = 1 -> 1:(X1Active' = X1Active);
   [e9_1] X1Active = 0 -> 1:(X1Active' = 1);
   [e10_1] Z1Active = 0 -> 1:(Z1Active' = Z1Active);
   [e11_1] Z1Active = 1 -> 1:(Z1Active' = Z1Active);
   [e12_1] X1Inactive = 1 -> 1:(X1Inactive' = 0);
endmodule

module GeneExpression
   Gene1 : [0..1] init 1;    Protein1 : [0..1] init 0;

   [e13_1] Gene1 = 1 & Protein1 = 0 -> 1:(Gene1' = 0) & (Protein1' = 1);
   [e14_1] Protein1 = 1 -> 1:(Protein1' = 0);
endmodule
```

We will treat these modules as *generic*, that is, we instantiate them (strictly, duplicate and rename in PRISM) for multiple occurrences. For brevity, we adopt the following convention. For generic module M, M_i denotes an *instance* of M with every variable and label renamed by an indexed form. For example, variable v becomes v_1 in module M_1.

3.2 Pathways

A pathway is a parallel composition of instances of the generic modules, renaming labels to coordinate synchronisation within the pathway.

Definition 1. *Let G be a set of generic modules. A pathway P has the form $f_1 X_1 \ |[L_1]| \ \ldots \ |[L_{n-1}]| \ f_n X_n$ where $X_1 \ldots X_n$ are instances of modules in G, $f_1 \ldots f_n$ are compositions of renamings and hidings and $L_1 \ldots L_{n-1}$ are labels.*

As an example, consider the expression of a pathway comprising the active receptor catalysing the activation of protein X and active protein Z catalysing the expression of Gene. This is defined by:

$$P_1 \ = \ Receptor_1 \ / \ \{i1_1\} \ \{e1_1 \ \leftarrow \ e5_1\} \ |[e5_1]| \ Cascade3_1 \ / \ \{i2_1\} \ \{e11_1 \ \leftarrow \ e13_1\}$$
$$|[e13_1]| \ GeneExpression_1$$

The $Receptor_1$ and $Cascade3_1$ modules synchronise on $e5_1$, and $Cascade3_1$ and $GeneExpression_1$ synchronise on $e13_1$; in both cases these are the only external labels that occur in both modules. Internal reactions $i1_1$ and $i2_1$ are hidden using the $/$ operator.

3.3 Composition of Pathways

In a similar way, pathways can be composed, synchronising on external labels.

Definition 2. *Given two pathways P_1 and P_2, with sets of external labels $ext(P_1)$ and $ext(P_2)$ resp., if $ext(P_1) \cap ext(P_2) = \{\}$ then the pathways are independent, otherwise there is crosstalk.*

As an example of independent pathways, consider P_1 and P_2 (shown in Figure 4), where

$$P_2 \ = \ Receptor_2 \ / \ \{i1_2\} \ \{e1_2 \ \leftarrow \ e5_2\} \ |[e5_2]| \ Cascade3_2 \ / \ \{i2_2\} \ \{e11_2 \ \leftarrow \ e13_2\}$$
$$|[e13_2]| \ GeneExpression_2$$

We synchronise P_1 and P_2 over the unused external reactions, i.e. the reactions that are otherwise never involved in synchronisation. In this case, $P_1 \ |[U]| \ P_2$ where $U = \{e2_1, e3_1, e4_1, e7_1, e8_1, e9_1, e10_1, e12_1, e14_1, e2_2, e3_2, e4_2, e7_2, e8_2, e9_2, e10_2, e12_2, e14_2\}$. Since the labels in U occur in only one pathway they cannot synchronise and their corresponding actions never execute.

P_1 and P_2 are independent pathways, since $ext(P_1) \cap ext(P_2) = \{\}$. We can also demonstrate independence through CSL properties. For example, the following CSL property expresses that it is not possible to activate P_1 without activating receptor R_1 or activate P_2 without activating receptor R_2.

$$P_1 \ |[U]| \ P_2 \ \models \ \mathbf{P}_{\leq 0} \ [\ F(RActive_1 \ = \ 0 \ \wedge \ Protein_1 \ = \ 1) \ \vee \ F(RActive_2 \ = \ 0 \ \wedge \ Protein_2 \ = \ 1) \ ]$$

In the next section we consider the case where the intersection is not empty; specifically, we consider examples of the five possible types of cross-talk between P_1 and P_2.

4 Cross-talk models

We now consider the five possible types of cross-talk between P_1 and P_2. In the previous section we synchronised P_1 and P_2 over U, the set of unused reactions. We now synchronise over E and U, $P_1 \ |[E, U]| \ P_2$, where $E = ext(P_1) \cap ext(P_2)$, i.e. the common external labels between P_1 and P_2, which may involve renaming. For each type we give a general description and then an explicit example.
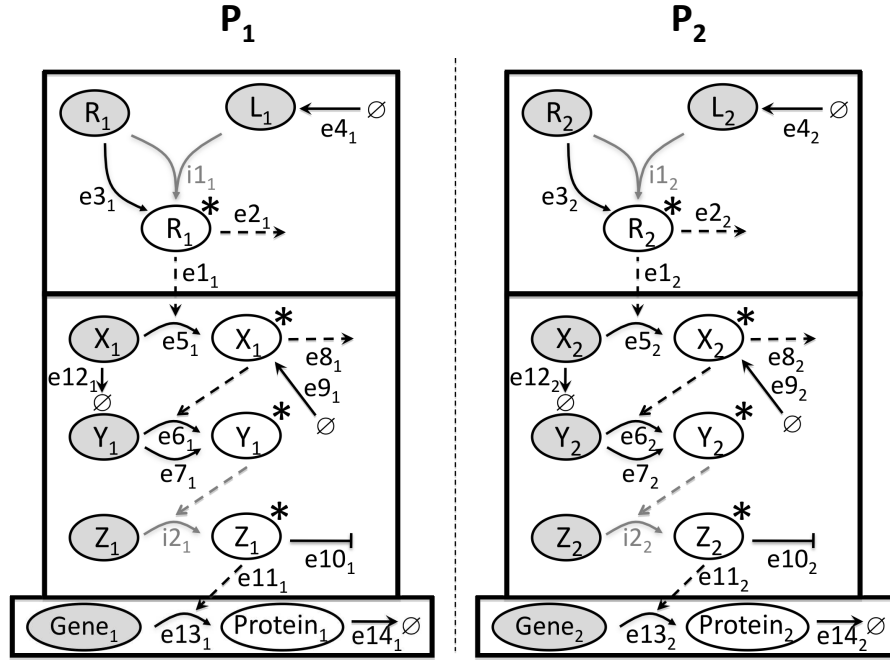
Figure 4: The two pathways P_1 and P_2 each comprising three instances of the generic pathway modules receptor, 3-stage cascade and gene expression. External reactions are denoted by black arcs, internal reactions by gray arcs and species that are present in the initial state by shaded ellipses.

We say that there is competition for a protein if the both pathways modify the protein such that they become unavailable to the other pathway. An enzyme that is shared between two pathways could be considered competition, however the duration of the enzyme-substrate complex is extremely short. As such, we ignore this step in our approach.

Signal Flow Cross-talk $P_1 \,|[E,\,U]|\, P_2$ has signal flow cross-talk if E only contains labels from two cascade modules or labels from a cascade module and catalysis/inhibition labels from a receptor module, in either case not setting up competition for a protein between the pathways. Example: provide an additional route to the activation of Y_1 through the X_2* enzyme. Synchronise $e8_2$, the enzymatic activity of X_2*, with $e7_1$ the alternative route to activate Y_1, by renaming $e8_2$ to $e7_1$ and synchronising between pathways on $e7_1$.

    $P_1 \,|[e7_1,\,U]|\, P_2\,\{e8_2 \leftarrow e7_1\}$
    where $U = \{e2_1,\, e3_1,\, e4_1,\, e8_1,\, e9_1,\, e10_1,\, e12_1,\, e14_1,\, e2_2,\, e3_2,\, e4_2,\, e7_2,\, e9_2,\, e10_2,\, e12_2,\, e14_2\}$.

Substrate Availability Cross-talk $P_1 \,|[E,\,U]|\, P_2$ has substrate availability cross-talk if E only contains labels from cascade modules or labels from a cascade module and catalysis/inhibition labels from a receptor module, in either case setting up competition for a protein between the pathways. Example: make both pathways compete for the activation of protein X_1 catalysed by their respective receptors $R1*$ and $R2*$. Within P_2, synchronise $e2_2$, the enzymatic activity of R_2*, with $e9_2$, the production of X_2*. Between pathways, synchronise $e12_1$, the degradation of X_1, with the $e9_2$. This produces a new reaction,

$X_1 \rightarrow X_2*$ with R_2* as the enzyme. Also, synchronise between pathways on $e5_2$ to block the following reaction, $X_2 \rightarrow X_2*$ with R_2* as the enzyme.

$P_1 \{e12_1 \leftarrow e9_2\} \|[e9_2, U]\| P_2'$

where $U = \{e2_1, e3_1, e4_1, e7_1, e8_1, e9_1, e10_1, e14_1, e3_2, e4_2, e5_2, e7_2, e8_2, e10_2, e12_2, e14_2]\|\}$

and $P_2' = Receptor_2 / \{i1_2\} \{e1_2 \leftarrow e5_2, e2_2 \leftarrow e9_2\} \|[e5_2, e9_2,]\| Cascade3_2 / \{i2_2\}$

$\{e11_2 \leftarrow e13_2\} \|[e13_2]\| GeneExpression_2.$

Receptor Function Cross-talk $P_1 \|[E, U]\| P_2$ has receptor function cross-talk if E only contains labels from a receptor module and catalysis/inhibition labels from either a receptor module or a cascade module. Example: provide an alternative route to activate receptor R_2 by the enzyme X_1*. Synchronise $e3_2$, the alternative route to active R_2, with $e8_1$, the enzymatic activity of X_1*.

$P_1 \{e8_1 \leftarrow e3_2\} \|[e3_2, U]\| P_2$

where $U = \{e2_1, e3_1, e4_1, e7_1, e9_1, e10_1, e12_1, e14_1, e2_2, e4_2, e7_2, e8_2, e9_2, e10_2, e12_2, e14_2\}.$

Gene Expression Cross-talk $P_1 \|[E, U]\| P_2$ has gene expression cross-talk if E only contains labels from a gene expression module and catalysis/inhibition labels from a cascade module. Example: connect the inhibiting activity of Z_2* to the expression of $Gene_1$. Synchronise $e10_2$, the inhibiting activity of Z_2*, with $e13_1$, the expression of $Gene_1$.

$P_1 \|[e13_1, U]\| P_2 \{e10_2 \leftarrow e13_1\}$

where $U =$

$\{e2_1, e3_1, e4_1, e7_1, e8_1, e9_1, e10_1, e12_1, e14_1, e2_2, e3_2, e4_2, e7_2, e8_2, e9_2, e12_2, e14_2\}.$

Intracellular Communication Cross-talk $P_1 \|[E, U]\| P_2$ has intracellular communication cross-talk if E only contains a protein degradation label from a gene expression module and a ligand production label from a receptor module. Example connect the degradation of $Protein_1$ to the production of L_2. Synchronise $e14_1$, the degradation of $Protein_1$, with $e4_2$, the production of L_2.

$P_1 \{e14_1 \leftarrow e4_2\} \|[e4_2, U]\| P_2$

where $U = \{e2_1, e3_1, e4_1, e7_1, e8_1, e9_1, e10_1, e12_1, e2_2, e3_2, e7_2, e8_2, e9_2, e10_2, e12_2, e14_2\}.$

4.1 Detecting Cross-talk

We perform Continuous Stochastic Logic (CSL) model checking in PRISM to detect the presence of cross-talk. For each of the 5 cross-talk models we compute the probability of three properties and note that cross-talk is detected by a change in probability compared with the independent model. The three properties are as follows.

Competitive Signal Flow (P_1): probability of signal flow through P_1 before P_2

$\mathbf{P}_{=?} \left[F(Protein_1 = 1 \wedge Protein_2 = 0) \right]$

Time-dependent Signal Flow (P_1): probability of signal flow through P_1 within 3 time units

$\mathbf{P}_{=?} \left[F_{\leq 3}(Protein_1 = 1) \right]$

Time-dependent Signal Flow (P_2): probability of signal flow through P_2 within 3 time units

$\mathbf{P}_{=?} \left[F_{\leq 3}(Protein_2 = 1) \right]$

	Competitive Signal Flow (P_1)	Time-dependent Signal Flow (P_1)	Time-dependent Signal Flow (P_2)
Independent Pathways	0.500	0.184	0.18473
Signal Flow Cross-talk (flow from P_2 to P_1)	0.638	0.304	0.18473
Substrate Availability Cross-talk (P_1 and P_2 compete for a protein)	0.500	0.141	0.14125
Receptor Function Cross-talk (P_1 activates P_2's receptor)	0.487	0.184	0.19257
Gene Expression Cross-talk (P_2 inhibits P_1's gene expression)	0.363	0.147	0.18473
Intracellular Communication Cross-talk (P_1 expresses P_2's ligand)	0.500	0.184	0.18477

The competitive signal flow property detects the presence of 3 of the 5 types of cross-talk. The probability of this property is significantly greater with signal flow cross-talk, due to the extra signal flow from P_2, and significantly less with gene expression cross-talk, due to inhibition of gene expression by P_2. There is a decrease in the probability of this property with receptor function cross-talk, however the decrease is small due to the likelihood that the cross-talk is initiated after P_2's receptor becomes activated. The competitive signal flow property does not detect the substrate availability cross-talk, as the cross-talk has an identical effect on each pathway. Intracellular communication cross-talk is also not detected, as the cross-talk occurs after *protein*$_1$ is expressed (the property does not refer to behaviour past this point).

To detect substrate availability and intracellular communication cross-talk we need to check the probability of independent signal flow through a pathway (hence, not in relation to another pathway). This is accomplished by checking the probability of signal flow through a pathway within a time bound, in this case (arbitrarily) chosen to be 3 time units. Substrate availability cross-talk has the effect of equally decreasing the time-dependent signal flow through both P_1 and P_2 due to the competition for a limited protein. Intracellular communication cross-talk has the effect of increasing only the signal flow through P_2, however the effect is marginal (5th decimal place). It is interesting to note that in Section 2.1.5 we identified intracellular communication cross-talk as a source of contention amongst the community and it proves difficult to detect in our analysis.

4.2   Characterising Cross-talk

We now define 5 CSL properties, each of which characterises a type of cross-talk and thus holds only in the respective model defined previously. The properties are simple liveness or safety properties and do not exploit the rate information in the model. These properties could equally be written in Computational Tree Logic, replacing the probabilistic operator with the universal (A) and existential (E) operators as appropriate.

Signal Flow Cross-talk (flow from P_2 to P_1)   it is possible to activate P_1 without activating receptor R_1
$\quad$ $\mathbf{P}_{>0} \left[F(RActive_1 = 0 \wedge Protein_1 = 1) \right]$

Substrate Availability Cross-talk (P_1 and P_2 compete for a protein)   it is not possible to activate both P_1 and P_2 (i.e. the pathways compete for a limited protein)

$$\mathbf{P}_{\leq 0} \left[F(Protein_1 = 1 \land Protein_2 = 1) \right]$$

Receptor Function Cross-talk (P_1 activates P_2's receptor)   it is possible to activate the receptor R_2 without using the ligand L_2

$$\mathbf{P}_{>0} \left[F(RActive_2 = 1 \land L_2 = 1) \right]$$

Gene Expression Cross-talk (P_2 inhibits P_1's gene expression)   it is not possible to express *protein*$_1$ if the signal has already passed through P_2

$$\mathbf{P}_{\leq 0} \left[F(Protein_1 = 1) \{Protein_1 = 0 \land Protein_2 = 1\} \right]$$

Intracellular Communication Cross-talk (P_1 expresses P_2's ligand)   it is possible to use and replenish ligand L_2

$$\mathbf{P}_{>0} \left[(L_2 = 1) \land (L_2 = 1) \, U \, ((L_2 = 0) \land (L_2 = 0) \, U \, (L_2 = 1)) \right]$$

5   Case Study: TGF-β/BMP, WNT and MAPK pathways

We applied our approach to a prominent biological case study of the cross-talk between the TGF-β/BMP, WNT and MAPK pathways. Details are taken from [5] and from discussions with a domain expert [14]. We consider the behaviour of the independent pathways and the various types of cross-talk. We note that the effects of cross-talk are not discussed in [5].

Our model of the pathways and their cross-talk is shown in Figure 5. To apply our module approach we need to expand our set of modules to: receptor, protein activation, 2-stage cascade, 3-stage cascade, translocation, protein binding and gene expression. This is a natural extension of our approach, and the extra modules act in a similar manner to the modules that have been discussed. Using the formal descriptions of cross-talk (Section 4), we have identified three types of cross-talk in this model.

We measure the output of the TGF-β/BMP pathway by the activity of the expression of Genes to Proteins. We use the following properties to compare the effects of cross-talk: ψ_1, the eventual expression of Genes, and ψ_2, the time-dependent expression of Genes (within 5 time units).

$$\psi_1 = \mathbf{P}_{=?} \left[F(Proteins = 1) \right], \qquad \psi_2 = \mathbf{P}_{=?} \left[F_{\leq 5}(Proteins = 1) \right]$$

Independent Pathways   With independence, the activation of the TGF-β/BMP pathway leads to gene expression within 5 time units, ψ_2, with probability 0.47 and eventual gene expression, ψ_1, with probability < 1 due to the inactivation of the receptor.

TGF-β/BMP and MAPK Cross-talk   There are two types of cross-talk between the TGF-β/BMP and MAPK pathways. *Signal flow: MAPK*$*$ proteins slow signal flow through the TGF-β/BMP pathway by deactivating the *R-Smads* and degrading *Smad*4. *Gene expression:* the *TF*$*$ and *AKT*$*$ proteins upregulate gene expression in the TGF-β/BMP pathway. Note that the appearance of the *AKT* and *PI*3*K* proteins in the MAPK pathway indicates an implicit cross-talk with the *AKT* and *PI*3*K* pathways respectively. The inclusion of cross-talk with the MAPK pathway can both provide alternative gene expression routes and block the TGF-β/BMP route, overall causing the probability of gene expression within 5 time units, ψ_2, to marginally decrease to 0.73. The probability of eventual gene expression, ψ_1, is 1 due to the consistent routes through the MAPK pathway.
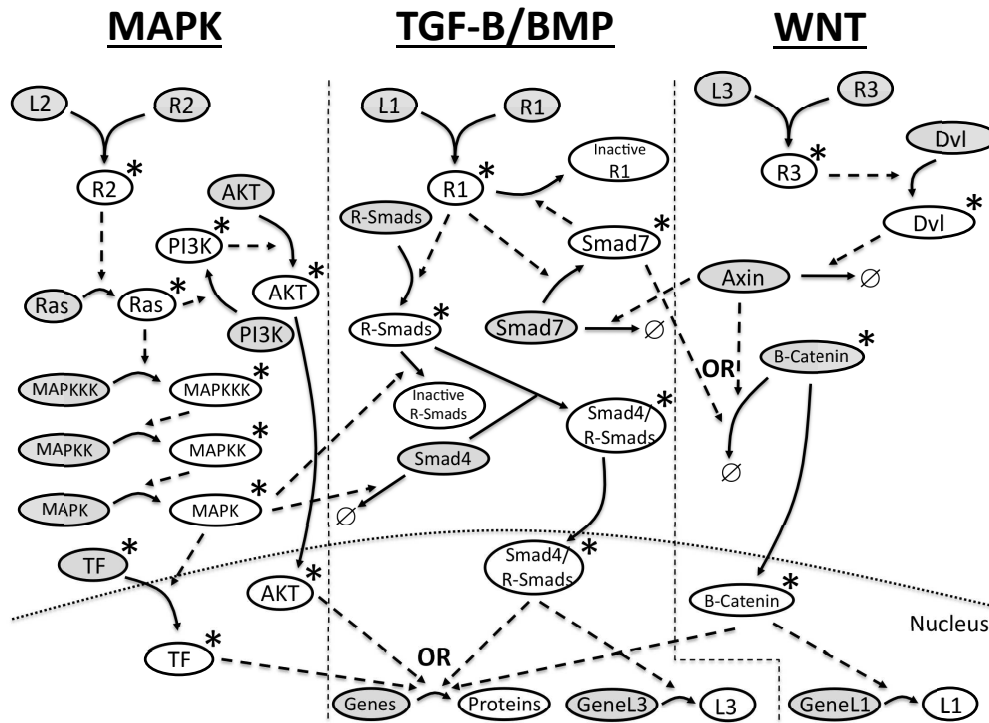
Figure 5: Cross-talk between the TGF-β/BMP, WNT and MAPK pathways. Species that are present in the initial state are denoted by shaded ellipses.

TGF-β/BMP and WNT Cross-talk   There are three types of cross-talk between the TGF-β/BMP and WNT pathways. *Signal flow:* the *Smad*7∗ protein degrades β-*Catenin* and the *axin* protein degrades *Smad*7[2]. *Gene expression:* the β-*Catenin* protein upregulates gene expression in the TGF-β/BMP pathway. *Intracellular communication:* the WNT pathway can cause the production of a ligand for the TGF-β/BMP pathway, and vice-versa. The inclusion of cross-talk with the WNT pathway can both provide an alternative route to gene expression and inhibit *Smad*7 which can inactivate the receptor for the TGF-β/BMP pathway. Overall this causes the probability of gene expression within 5 time units, ψ_2, to marginally increase to 0.76. The probability of eventual gene expression, ψ_1, is still < 1 due to the degradation of the β-*Catenin* protein.

TGF-β/BMP, WNT and MAPK Cross-talk   The cross-talk between all three pathways is the union of the two cross-talk scenarios above. The effect of both WNT and MAPK cross-talk to the TGF-β/BMP pathway is additive. The probability of ψ_2 has risen to 0.88, compared with the single cross-talks of WNT and MAPK with probability 0.76 and 0.73 respectively. The inclusion of the MAPK cross-talk provides consistent routes to gene expression and hence the probability of ψ_1 is 1.

---

[2]this cross-talk is discussed further in Section 6

6 Discussion

Categorisation   The categorisation of signal flow cross-talk between the WNT and TGF-β/BMP pathways in the previous section may be open to discussion. In Figure 5 *Axin* degrades *Smad*7. Rather than being part of the signal flow from the TGF-β/BMP receptor to gene expression, *Smad*7 actually deactivates the TGF-β/BMP receptor. One could consider this cross-talk as receptor function cross-talk with an intermediate (*Smad*7). Our approach categorises this as signal flow because the set of labels involved is an catalysis label from a protein module and a degradation label from a protein activation module.

Quantitative Detail   We have demonstrated our approach in this paper on models with a low level of quantitative detail. As such, the probability values resulting from CSL model checking can only be used to compare between models. However, with more quantitative detail, further interpretation of our analysis results would be possible. For example, the properties concerning the probability of time-dependent gene expression between cross-talk models would become a meaningful assessment of the strength of the cross-talk.

Feature Interaction   There may be an interesting analogy with feature interactions in telecommunications and software systems. Features, or services, in these systems are additional functionality (additional to the core). They are often added incrementally, by various developers, at various times (e.g. due to deregulation). A possible consequence is interactions between the new features themselves, or with the core system, causing some features or the core to behave in new, sometimes undesirable ways. An open question is whether techniques developed to model and detect features and interactions may be applicable to pathway cross-talk. Moreover, a common problem is lack of universal definition of pathway/feature; it would be interesting to investigate if concepts such as the feature construct of [15] would be useful in the pathway paradigm. We note that many approaches to interaction detection are based on temporal logic descriptions of behaviour.

7 Related Work

The literature contains a limited number of applications of computational techniques to the study of signalling pathway cross-talk.

There are several examples whereby a proposed method of cross-talk between two pathways is expressed and analysed as a computational model. For example [16] and [17] use Ordinary Differential Equations (ODEs) to model the cross-talk between the MAPK pathway with the AKT and PKC pathways respectively. The work in [18] is similar, however the modelling technique used to model the cross-talk between the growth factor pathway and the Integrin pathway is stochastic networks. Furthermore, a more formal notation of Petri nets has been used in [19] to model the cross-talk between the pathways involved in Apoptosis decision-making.

The computational analysis of cross-talk models have produced some interesting results in the literature. For example, [7] has produced a model containing substrate availability cross-talk between the hyperosmolar and the pheromone MAPK pathways. The question the authors try to answer is how these pathways maintain signal specificity given that they share common proteins. The authors analyse two models which can account for the signal specificity, one which contains mutual inhibition between pathways to limit signal bleed-through and one which contains scaffold proteins. In [17] the computational analysis shows that cross-talk has an effect on whether the model exhibits bistability. With pathway

cross-talk, the behaviour of key proteins switches from transient to sustained activation upon varying the duration of the signal.

We have found only one paper on the application of formal models to cross-talk, [20], which contains a model of the multiple modes of cross-talk between the EGFR and LIN-12/Notch signalling pathways. A discrete, dynamic, state-based model is developed in using the language of Reactive Modules. Model checking is used to check the validity of the model and to generate "new biological insights into the regulatory network governing the cell fate". However, this work differs from ours because it concerns intercellular cross-talk within a multi-cellular model. The application of the term cross-talk to intercellular communication is often considered a misnomer amongst the community, and hence we have focussed our analysis to intracellular cross-talk in a single-cell model.

8   Conclusion and Future Work

In this paper we have explored formal modelling and analysis techniques for signalling pathway cross-talk. We note that formal methods, whilst applied frequently in the literature to signalling pathways, have been largely ignored in the study of cross-talk. The aim of this paper has been to show that formal methods are both well-suited to and a natural choice for this area.

Our first contribution is a novel categorisation of cross-talk, drawing from examples in the literature.

The second contribution is the definition of a pathway and pathway composition, based on a set of generic pathway modules with internal and external reactions, and renaming and synchronisation operations. We can compose pathway independently, or with cross-talk. Furthermore, we also find that temporal logic descriptions of behaviours are suitable to detect and characterise cross-talk. The approach is illustrated with example pathways.

The third and final contribution is the application to cross-talk between the TGF-β/BMP pathway and two other pathways.

Several future directions have been identified. As suggested earlier in the paper, we wish to apply our approach to models with a higher level of quantitative detail to make predictions and generate insights about the biological effects of cross-talk. Furthermore, we wish to assess how the effect of cross-talks differs to that of standard pathway motifs. Hence, is there a reasonable alteration of the pathway model (e.g. addition of a feedback loop) that gives the same behaviour as a potential cross-talk? We also wish to assess how the effectiveness of pathway intervention techniques (e.g. drugs, gene knockouts) changes with the addition of cross-talk. Finally, a larger question is how the temporal ordering of signals affects the detectability and behaviour of cross-talk (i.e. do the pathways hold a "biochemical history" of signalling events?).

The PRISM models and CSL properties used in this paper can be found at:

`www.dcs.gla.ac.uk/~radonald/fbtc2010/.`

References

[1] Brightman, F., Fell, D.: Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. FEBS Lett 482 (2000) 169–174

[2] Taniguchi, C.M., Emanuelli, B., Kahn, R.C.: Critical nodes in signalling pathways: insights into insulin action. Nat Rev Mol Cell Biol **7**(2) (February 2006) 85–96

[3] Catt, I.: Crosstalk (Noise) in Digital Systems. Electronic Computers, IEEE Transactions on **EC-16**(6) (1967) 743–763

[4] Schwartz, M.A., Ginsberg, M.H.: Networks and crosstalk: integrin signalling spreads. Nat Cell Biol **4**(4) (April 2002) E65–E68

[5] Guo, X., Wang, X.F.: Signaling cross-talk between TGF-β/BMP and other pathways. Cell Research **19** (2009) 71–88

[6] NIH: "Insulin Signaling And Receptor Cross-Talk" Program Announcement PA-07-058. November 20, 2006.

[7] McClean, M.N.N., Mody, A., Broach, J.R.R., Ramanathan, S.: Cross-talk and decision making in MAP kinase pathways. Nat Genet (January 2007)

[8] Katzenellenbogen, B.: Estrogen receptors: bioactivities and interactions with cell signaling pathways. Biol. Reprod. **54** (1996) 287–293

[9] Bosscher, K.D., Berghe, W.V., Haegeman, G.: Cross-talk between nuclear receptors and nuclear factor κb. Oncogene **25** (2006) 6868–6886

[10] Orton, R.J., Sturm, O.E., Vyshemirsky, V., Calder, M., Gilbert, D.R., Kolch, W.: Computational modelling of the receptor tyrosine kinase activated MAPK pathway. Biochemical Journal **392**(2) (2005) 249–261

[11] PRISM Website: PRISM - Probabilistic Symbolic Model Checker. `http://www.prismmodelchecker.org`

[12] Calder, M., Vyshemirsky, V., Orton, R., Gilbert, D.: Analysis of Signalling Pathways using Continuous Time Markov Chains. Trans. on Computat. Syst. Biol. VI **4220** (2006) 44–67

[13] Ciochetta, F., Degasperi, A., Hillston, J., Calder., M.: Some investigations concerning the CTMC and the ODE model derived from Bio-PEPA. Proceedings of FBTC 2008, ENTCS **17350** (2008)

[14] Breitling, R.: Private correspondence

[15] Plath, M., Ryan, M.: The feature construct for SMV: Semantics. In Calder, M., Magill, E., eds.: Feature Interactions in Telecommunications and Software Systems VI, IOS Press (Amsterdam) (2000) 129–144

[16] Hatakeyama, M., Kimura, S., Naka, T., Kawasaki, T., Yumoto, N., Ichikawa, M., Kim, J., Saito, K., Saeki, M., Shirouzu, M., Yokoyama, S., Konagaya, A.: A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. Biochemical Journal **373 Pt. 2** (2003) 451–463

[17] Sreenath, S.N., Soebiyanto, R., Mesarovic, M.D., Wolkenhauer, O.: Coordination of crosstalk between MAPK-PKC pathways: an exploratory study. Systems Biology, IET **1** (2007) 33–40

[18] Bauer, A., Jackson, T., Jiang, Y., Rohlf, T.: Stochastic network model of receptor cross-talk explains pro- and anti-angiogenic control. under review (2009)

[19] Heiner, M., Koch, I., Will, J.: Model validation of biological pathways using Petri nets–demonstrated for apoptosis. Journal BioSystems **75** (2004) 15–28

[20] Fisher, J., Piterman, N., Hajnal, A., Henzinger, T.A.: Predictive Modeling of Signaling Crosstalk during C. elegans Vulval Development. PLoS Computational Biology **3**(5) (May 2007) e92+