

# **Model checking**

David Gilbert <u>david.gilbert@brunel.ac.uk</u>

http://people.brunel.ac.uk/~csstdrg/

Centre for Systems & Synthetic Biology and School of Information Science Computing and Mathematics Brunel University, UK

Special thanks to Robin Donaldson, University of Glasgow, UK

# **Overview**

- What is model checking
- What can we do with model checking
- Model checking for systems biology
- MC: qualitative, continuous & stochastic
- Analytical & simulative MC
- Simulative MC & biochemical pathways
- Advanced example: parameter fitting using model checking
- Model checking for synthetic biology
- MC systems illustrated: MC2, Marcie

# Model Checking

In a sentence:

 "Formally check whether a model of a biochemical system does what we want"

Components:

- A model
  - the current description of a biochemical system of interest
- A property
  - a property which we think the system should have
- A model checker
  - a program to test whether the model has the property

### Model Checking **Biochemical Pathways**



# What can we do with Model Checking?

- Model validation:
  - Show that your model of the pathway matches the lab data (...stochastic...)
- Model analysis:
  - In a collection of variants of a model (e.g., *in silico* gene knock-outs), which models show a certain behavior? (loss of oscillations...)
- Model development:
  - If the model doesn't do what we want, change the model automatically until it does! (parameters, structures,...)
- Model finding:
  - Many models in a database, can use model checking to query the database
  - "Give me all the models in the database which oscillate"
- Biosystem verification:
  - Does the constructed system do what we intended?

# **Describing experimental data**

- Biologists will often talk in qualitative or semiquantitative language (trends).
  - "this protein peaks after 5 minutes, then falls to half concentration"
  - Often quite certain about time,
  - But not about concentrations

### Lab data versus simulations



#### Simulation



#### Experiment

# **Properties...**

Examples:

- After 100 seconds the concentration of Protein1 is stable
- Protein1 peaks and falls
- Protein1 peaks and stays constant
- Protein1 peaks before Protein2
- Protein1 oscillates 4 times in 5,000 seconds
- Molecules of Protein2 are required for molecules of Protein1 to be created

# To formally express properties we use temporal logics

Various logics each with different expressivity

Branching logics consider all branching time lines.

- Computational Tree Logic (CTL)
- Continuous Stochastic Logic (CSL)

"There is a possibility that I will stay hungry forever."

"There is a possibility that eventually I am no longer hungry."

Linear time logics consider separately all single time lines.

– Linear-time Temporal Logic (LTL, LTLc, PLTLc)

"I am hungry."

"I am always hungry", "I will eventually be hungry", "I will be hungry until I eat something".

David Gilbert



Gilbert, Heiner and Lehrack. ``A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets." Proc CMSB 2007

David Gilbert

# Models

- Qualitative:
  - Petri nets,...
- Continuous:
  - ODEs, Continuous Petri nets
- Stochastic:
  - Stochastic process algebras, Stochastic Petri nets, P-systems, (Chemical master equations)...

### Levchenko Model



#### David Gilbert

# **CTL Computational Tree Logic**

- EX  $\phi$  : if there is a state reachable by one step where  $\phi$  holds.
- EF φ : if there is a path where φ holds finally,
   i.e., at some point.
- EG φ : if there is a path where φ holds globally,
   i.e., forever.
- E ( $\phi_1 U \phi_2$ ) : if there is a path where  $\phi_1$  holds until  $\phi_2$  holds

# **CTL Computational Tree Logic**

- AX φ : if φ holds for all states which are reachable by one step.
- AF φ : if φ holds finally (at some point) for all paths.
- AG φ : if φ holds globally (i.e. for ever) for all paths.
- A (φ<sub>1</sub> U φ<sub>2</sub>) : if φ<sub>1</sub> holds until φ<sub>2</sub> holds for all paths.

# Infinite run (beginning)

### Generated from the input-output T-invariant



# **Qualitative: CTL properties**

 property Q1: The signal sequence predicted by the partial order run of the I/O T-invariant is the only possible one. I.e., starting at the initial state, it is necessary to pass through states RafP, MEKP, MEKPP and ERKP in order to reach ERKPP.



# **Qualitative: CTL properties**

- property Q2: Dephosphorylation takes place independently. E.g., the duration of the phosphorylated state of ERK is independent of the duration of the phosphorylated states of MEK and Raf.
  - ( EF [ Raf ∧ (ERKP ∨ ERKPP )] ∧ EF [ RafP ∧ (ERKP ∨ ERKPP )] ∧ EF [ MEK ∧ (ERKP ∨ ERKPP )] ∧ EF [ (MEKP ∨ MEKPP ) ∧ (ERKP ∨ ERKPP )] )
- DSSZ-MC (BDD-CTL, IDD-CTL) Cottbus (qualitative: pn2009 paper)

# **CSL Continuous Stochastic Logic**

- Replaces the path quantifiers (E, A) in CTL by the probability operator P<sub>▷⊲p</sub>, where ▷⊲p specifies the probability p of the formula.
- $P_{=?}[X \phi]$ : prob there is a state reachable by one step where  $\phi$  holds.
- P<sub>=?</sub> [F φ] : prob there is a path where φ holds finally, i.e., at some point.
- P<sub>=?</sub> [G φ] : prob there is a path where φ holds globally, i.e., forever
- $P_{=?}$  [( $\varphi U \varphi$ )] : prob there is a path where  $\varphi$  holds until  $\varphi$  holds

# Levels



# **Stochastic: CSL properties**

 property S1: What is the probability of the concentration of RafP increasing, when starting in a state where the level is already at L (the latter side condition is specified by the filter given in braces)?



### $P_{=?}[(RafP = L)U^{<=100}(RafP > L){RafP = L}]$

# **CSL** properties

 property S2: What is the probability that, given the initial concentrations of RafP, MEKPP and ERKPP being zero, the concentration of RafP rises above some level L while the concentrations of MEKPP and ERKPP remain at zero, i.e. RafP is the first species to react?

 $P_{=?}[((MEKPP = 0) \land (ERKPP = 0))U^{<=100}] (RafP > L){(MEKPP = 0) \land (ERKPP = 0) \land (RafP = 0)}$ 



# LTL Linear Temporal Logic

- G ( $\phi$ ) :  $\phi$  always happens
- F ( $\phi$ ) :  $\phi$  happens at some time
- X ( $\phi$ ) :  $\phi$  happens in the next time point
- $\phi_1 \cup \phi_2$  :  $\phi_1$  happens until  $\phi_2$  happens

### Syntactic sugar (also for CSL)

•  $\phi_1 \{\phi_2\}$  :  $\phi_1$  happens from the first time  $\phi_2$  happens, where no temporal operators in  $\phi_2$ 

### **Time-series (continuous)**



# **Continuous: LTLc properties**

 property C1: The concentration of RafP rises to a significant level, while the concentrations of MEKPP and ERKPP remain close to zero; i.e. RafP is really the first species to react.





David Gilbert

# **Continuous: LTLc properties**

• **property C2**: if the concentration of RafP is at a significant concentration level and that of ERKPP is close to zero, then both species remain in these states until the concentration of MEKPP becomes significant; i.e. MEKPP is the second species to react.

( (RafP > 0.06) ∧ (ERKPP < 0.0002) )  $\Rightarrow$ 

( (RafP > 0.06) ∧ (ERKPP < 0.0002) ) U (MEKPP > 0.004)



David Gilbert

# **Continuous: LTLc properties**

 property C3: if the concentrations of RafP and MEKPP are significant, they remain so, until the concentration of ERKPP becomes significant; i.e. ERKPP is the third species to react.

( (RafP > 0.06) ∧ (MEKPP > 0.004) ) ⇒ ( (RafP > 0.06) ∧ (MEKPP > 0.004) ) U (ERKPP > 0.0005)



# Analytical vs Simulative Model Checking

### Analytical:

- Exact probabilities & prove properties
- A model state is an association of #molecules/levels to each of the species
  - Protein1 has 10 molecules & Protein2 has 20 molecules
- Analytical assesses every state that the model can be in (reachable states)
- State space can grow even worse than exponentially with increasing molecules, or even be infinite!
- Stochastic model checking with even as little as 12 molecules/levels can be impossible with today's technology

### Simulative:

- Instead of analysing the constructed state space, analyse simulation outputs
- Simulate the model X times and check these simulations
- Simulation run = finite path through the state space
- Can't prove probabilities

# **Probabilistic analytical model checkers**

### PRISM

- Specification language based on reactive modules language
- Supports 3 types of models
  - Discrete time Markov chains (DTMCs)
  - Continuous time Markov chains (CTMCs)
  - Markov decision processes (MDPs)
- Logic: PCTL or CSL
- Symbolic
- Communication via shared events
- Synchronous execution (apart from MDPs)

Quantitative analysis using costs/rewards Can run experiments

### IDD-CSL

New system under development from Cottbus (stochastic 1<sup>st</sup> prototype AWPN 2008, CMSB2009)

# Simulative Model Checking

- In-line: check the observations as they arrive
  - Requires complex computational machinery: 'combine' simulator & model checker
  - Good for biochemical observations
  - Don't always need to finish the experimental run
- **Off-line**: check the observations after all have been generated
  - Easier to implement computationally (simulate then check)
  - Need to always define when to 'stop' generating observations

### Simulation-based Model Checking Biochemical Pathways



# **MC2: Monte Carlo Model Checker**

- Offline Monte Carlo Model Checker for PLTLc properties.
- Operates on a finite set of simulations simulative approach
- Typically, many stochastic simulations to approximate probabilities
  - Approximate probability = fraction of simulations which satisfy the property over the #simulations
- Monte Carlo approximation 2 approximations made:
  - finite number of simulations
  - Simulations of finite length

# Monte Carlo Model Checker

- The set of time series data can be:
  - Set of stochastic runs
  - A single continuous run
  - A parameter scan
  - Lab data!
- We could use simulation output from;
  - ODE, SDE, CTMC, Gillespie, hybrid approaches, multicellular simulation, open models
- Or experimental data from the wet lab

### **MC2 with ODE Output**





David Gilbert

### **MC2 with Gillespie Output**





### **MC2 with Gillespie Output**

 $P_{=?}[F(X > 5)]$ 

=> P = 4/6



David Gilbert

# **MC2 versus Analytical Checkers**

- MC2
  - Probabilities are derived from fraction of #finite simulation outputs satisfying property
  - Provides an approximation of the probability because
    - Finite set of simulation outputs and simulation outputs are finite size
- PRISM
  - Exact probabilities are produced but much higher cost
  - Exact, need to construct the state-space
  - State-space is worse than exponential or even infinite in #levels and #protein types.
  - Gilbert *et al.* (2007) can only use up to 8 concentration levels in PRISM – with MC2 we can use easily 4,000.

Gilbert *et al.* (2007), "A unifying framework for modelling and analysing biochemical pathways using Petri nets". In Proc. CMSB 2007, pages 200-216.

David Gilbert

# **MC2 versus Analytical checkers**

- Can easily perform model checking on a cluster Fast!
- Fast, but it is an approximation...
- No need to *explicitly* impose time-bounds in formulae on operators for efficiency as done in Gilbert *et al.* (2007)
   P<sub>=?</sub>[ (RafP = X) U<sub><=100</sub> (RafP > X) { RafP = X } ]
  - but implicit time bounds due to finite simulation runs...

Gilbert *et al.* (2007), "A unifying framework for modelling and analysing biochemical pathways using Petri nets". In Proc. CMSB 2007, pages 200-216.

David Gilbert

# **Stochastic Analysis**

Check the property S2:

"What is the probability that RafP will reach concentration X while MEKPP and ERKPP remain at 0?"

 $P_{=?} [ (MEKPP = 0 ^ ERKPP = 0) U (RafP > X)$  ${RafP = 0 ^ MEKPP = 0 ^ ERKPP = 0} ]$ 



David Gilbert, Monika Heiner and Sebastian Lehrack (2007). A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets. Proc CMSB 2007

David Gilbert

# Checking the property at varying levels



Monika Heiner, David Gilbert, and Robin Donaldson (2008), Petri Nets for Systems and Synthetic Biology. In M Bernardo, P Degano, and G Zavattaro (Eds.): Formal Methods for Systems Biology SFM 2008, Springer LNCS 5016 David Gilbert Model checking

# PLTLc language, specifics

(1) LTLc [Fages et al.] - extension of LTL with numerical constraints.

(2) PLTLc [Donaldson et al]: Add probabilistic operator & free variables

E.g.

Free variable \$X always greater than the concentration of Protein.
 P<sub>=?</sub> [ G( \$X > [Protein] ) ]

Model checking of PLTLc properties returns:

- Probability of behaviour: some value is always greater than concentration of protein
- The domain of free variable \$X for which the behaviour holds true

Fages et al. "On the Analysis of Numerical Data Time Series in Temporal Logic." In *Proc. CMSB2007, pp48–63. LNCS/ LNBI4695, Springer, 2007.* 

Donaldson and Gilbert. A Model Checking Approach to the Parameter Estimation of Biochemical Pathways In proceedings CMSB 2008 (Computational Methods in Systems Biology). LNCS 5307/2008, pp269-287

David Gilbert

# **PLTLc language**

- Continuous model with [Protein] between 0...8 in simulation output
- P<sub>=?</sub> [ G( \$X > [Protein] ) ]
- Probability =1 : there is a value always greater than [Protein]
- Domain of \$X is 9...inf : these are the values always greater than [Protein] :

• Probabilistic language, interpret this as a probabilistic domain:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	 $\infty$
0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

# **PLTLc language**

- Stochastic model has varied behaviour, so probabilistic domain ranges 0...1.
- Could look something like:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	•••	α	)
0	0	0	0	0	0	0	0	0	0.1	0.2	0.67	0.78	81	••••	1	

# Qualitative to quantitative descriptions in PLTL

- From qualitative to quantitative
- Qualitative:

Protein rises then falls
P=? [ ( d[Protein] > 0 ) U ( G( d[Protein] < 0 ) ) ]</pre>

### Semi-qualitative:

Protein rises then falls to less than 50% of peak concentration P=? [(d[Protein] > 0) U(G(d[Protein] < 0  $\land$  F([Protein] < 0.5 \* max[Protein]))]

#### Semi-quantitative:

Protein rises then falls to less than 50% of peak concentration by 60 minutes P=? [(d[Protein] > 0) U(G(d[Protein] < 0)  $\land$  F(time = 60  $\land$  Protein < 0.5 \* max[Protein]))]

#### Quantitative:

Protein rises then falls to less than <u>100µMol</u> by 60 minutes P=? [ ( d[Protein] > 0 ) U ( G( d[Protein] < 0 )  $\land$  F ( time = 60  $\land$  Protein < 100 ) ) ]

# Parameter estimation using the Monte Carlo Model Checker

*Continuous* Brightman & Fell model:

- The EGF signal transduction pathway produces transient Ras, MEK and ERK activation whereas NGF stimulation produces sustained activation.
- Parameter V28 has the the highest probability of generating the desired behaviour, but requires 40-fold increase in value



15

Brightman & Fell, FEBS Lett 2000. "Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells"

David Gilbert

# **Response with EGF vs. NGF signal**



# **Desired Behaviour in PLTLc**

The desired (sustained) NGF behaviour of the pathway was written informally in the original paper.

We can formalise it in PLTLc as:

```
Sustained Ras: Active Ras peaks within 2 min to max 20% total Ras and stable between 5%..10%
```

$$\begin{split} \mathsf{P}_{=?} \left[ \begin{array}{c} \mathsf{d}(\mathsf{active Ras}) > 0 \quad \mathsf{U} \ ( \ \mathsf{time} \leq 2 \ \land \ \mathsf{active Ras} \geq 0.15 \ast \mathsf{total Ras} \\ \land \ \mathsf{active Ras} \leq 0.2 \ast \mathsf{total Ras} \ \land \ ( \ \mathsf{d}(\mathsf{active Ras}) < 0) \\ \mathsf{U} \ ( \ \mathsf{G}( \ \mathsf{active Ras} \geq 0.05 \ast \mathsf{total Ras} \ \land \ \mathsf{active Ras} \leq 0.10 \ast \mathsf{total Ras} \ ) \ ) \ ) \end{array} \right] \end{split}$$

Sustained MEK: Active MEK peaks in 2 to 5 min and is stable between 40%..50% of peak value

 $P_{=?} [ d(MEKPP) > 0 U ( time \ge 2 \land time \le 5 \land d(MEKPP) < 0 U ( G(MEKPP \ge 0.40*max(MEKPP) \land MEKPP \le 0.50*max(MEKPP) )) ) ]$ 

Sustained ERK: Active ERK peaks in 2 to 5 min and is stable between 85%..100% of peak value  $P_{=?} [ ( d(ERKPP) > 0 ) U ( time \ge 2 \land time \le 5 \land d(ERKPP) < 0 U ( G(ERKPP \ge 0.85 * max(ERKPP) ) ) ) ]$ 

Robin Donaldson and David Gilbert (2008). A Model Checking Approach to the Parameter Estimation of Biochemical Pathways In proceedings CMSB 2008 (Computational Methods in Systems Biology). LNCS 5307/2008, pp269-287

David Gilbert

# **Critical parameters**

Critical parameters can produce sustained activation of Ras, MEK or ERK.

Used to give an idea which parameters to vary

Method:

 Vary the kinetic rate constant parameters in range ± 2 orders of magnitude from original value.

- Perform 1,000 simulations using different values for each parameter, linearly spaced in the range

- The 'significance values' are the fraction of values in the range which give rise to sustained behaviour for each protein

- Found through model checking

sustained sustained sustained Ras MEK ERK parameter V\_20 0.010.00.001V\_24 0.076 0.00.0 $V_25$ 0.0230.0010.0 V\_27 0.614 0.00.0V\_28 0.4780.1510.679k1\_14 0.0 0.7780.0k1\_16 0.00.00.001k1\_18 0.8070.0010.0k2\_18 0.1910.00.0 $Km_20$ 0.0010.00.797kcat\_21 0.0010.00.688 kcat\_23 0.0010.00.186Km\_23 0.1210.00.0 $Km_25$ 0.0010.1570.0kcat\_26 0.0010.0 0.0Km\_26 0.00.00.005

### 2000 models, 100 generations, 2.10<sup>5</sup> simulations/checks Model construction



NGF (sustained) model by varying the values of the *critical* kinetic rate constants.

I.e. minimise the distance of the model to sustained behaviour.

Could vary the initial concentrations, or topologies (ongoing research).

Genetic algorithm:

- Define a parameter space (±2 orders of magnitude for each parameter)
- Initialise a population of models randomly throughout parameter space
- For each generation of the algorithm:
  - Perform genetic operations on binary representations of the models' parameter values (crossover, mutation, reproduction)
  - Evaluate all models' fitness values using model checking fitness is the distance to sustained behaviour
  - Probabilistically select models to survive to the next generation based on their fitness value



# **Fitness function using PLTLc**

#### **Probability**:

- Can optimise the probability of a behaviour, which works fine on stochastic models.
  - On continuous models, the probability is boolean so not good in a fitness function – no gradient

#### Free variables:

- Can use the free variables in a PLTLc behaviour, works for continuous or stochastic models.
  - Can always get a numerical value for the fitness function, even in continuous models – good for search algorithm
- We specify the behaviour in PLTLc and at the same time characterise the 'tail' of the peak in a free variable.
- We have an idea of the desired behaviour of the tail and can calculate the distance, using the free variables, to give us a numerical value for the fitness function, whilst....
- the behaviour in PLTLc enforces a peak at the right position.

David Gilbert

# **Fitness function using PLTLc**

Fitness function for sustained ERK

Enforce a peak between time 2 and 5 and characterise the tail in \$ErkppTail:  $P=? [(d(ERKPP) > 0) U(time \ge 2 \land time \le 5 \land d(ERKPP) < 0 \land G(ERKPP \ge $ErkppTail))]$ 

• Fitness function is distance between tail and 85% of the peak height (sustained activation), calculated using the probabilistic domains:



### **Distance metric**

Uses Residual Sum of Squares Over probabilistic domains in range m,n

$$RSS(X,X',m,n) = \sum_{i=m}^{n} |\$X(i)-\$X'(i)|^{2}$$

Dist(Model,Model')= (dist(Ras)+dist(MEKpp)+dist(ERKpp))/3

# Results

- Built a fitness function for sustained Ras, MEK and ERK
- Ran the genetic algorithm with 100 generations with results:



- Original model of the NGF signalling pathway varying V28 (dotted)
- Best model returned when varying the critical parameters (solid)
- Critical parameters without V28 (dashed).

The best model returned when varying the critical parameters only required a **16-fold** increase in V28 (compared with 40-fold in original paper)

#### Even possible to get similar behaviour without varying V28

Robin Donaldson and David Gilbert (2008). A Model Checking Approach to the Parameter Estimation of Biochemical Pathways In proceedings CMSB 2008 (Computational Methods in Systems Biology). LNCS 5307/2008, pp269-287 David Gilbert Model checking

# **Related work**

- Francois Fages, Biocham.
  - On a Continuous Degree of Satisfaction of Temporal Logic Formulae with Applications to Systems Biology Aurélien Rizk, Grégory Batt, François Fages and Syvain Soliman. Comptational Methods in Systems Biology CMSB'08

# Model searching

- Query the database on the behaviour of the models
- SIMAP project has a database of MAPK pathway models
- We are using MC2(PLTLc) as a SQL for models of biochemical systems
- Also want to use it as a SQL for lab data
- For example, these questions
  - "What are the models where ERK oscillates?"
  - "Under what conditions does ERK not behave as we observe?"
  - "What are the behaviours of the pathway not backed up by lab data?"

# **Model searching**

#### Peaks at least once

# (rises then falls below 50% max concentration)

- P<sub>>=1</sub>[ ErkPP <= 0.50\*max(ErkPP) ∧ d(ErkPP) > 0 U ( ErkPP = max(ErkPP) ∧ F( ErkPP <= 0.50\*max (ErkPP) ) ) ]
- Brown
- Kholodenko
- Schoeberl

#### **Rises and remains constant**

(99% max concentration) P<sub>>=1</sub>[ErkPP <= 0.50\*max(ErkPP) ∧ (d(ErkPP) > 0) U (G (ErkPP >= 0.99\*max(ErkPP))) ]

• Levchenko

#### **Oscillates at least 4 times**

 $P_{>=1}[ F(d(ErkPP) > 0 \land F(d(ErkPP) < 0 \land ...)) ]$ 

Kholodenko

#### David Gilbert





### **Database Web Interface**



MAPK Model Database: Originally designed by Richard Orton

developed and maintained by Qian (Pam) Gao

### MAPK



David Gilbert

# Models – Schoeberl Model (2000)

	MAPK Mo	del Database
MAPK	Home   MAPK	Models   Search   Links
Menu	Home > Mod	els > <u>Schoeberl2002</u>
PubMed	Schoeberl20	02 Model Information
Journal	author	Schoeberl2002
	title	Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF
PDF SBML	description	Ve present a computational model that offers an integrated quantitative, dynamic, and topological representation of
<u>Output</u> <u>Species</u> <u>Reactions</u>		Intracellular signal networks, based on known components of epidermal growth factor (EGF) receptor signal pathways. The model provides insight into signal-response relationships between the binding of EGF to its receptor at the cell surface and the activation of downstream proteins in the signaling cascade. It shows that EGF-induced responses are remarkably stable over a 100-fold range of ligand concentration and that the critical parameter in determining signal efficacy is the initial velocity of receptor activation. The predictions of the model agree well with experimental analysis of the effect of EGF on two downstream responses, phosphorylation of ERK-1/2 and expression of the target
	anaciaa	gene, c-fos.
	reactions	148
	picture	
	picture	Ras Ras Ras Ras Ras Ras Plase1 Welk Plase2 Plase2 Plase2 Plase2 Plase2 Plase2 Plase2 Plase3 Pl
	simulation	Schoebert2002b Simulation Plots
David Gilbert		Model checking

# Validation & verification

- Validation 'You built the right product?'.
  - Product / system accomplishes its intended requirements.
  - Model / simulation are accurate representations of the real world
- Verification 'You built the product right?'.
  - System complies with its specification
  - Model / simulation accurately represent the specifications

# Synthetic biology development cycle



David Gilbert

### **Model checking for Synthetic Biology**



# Acknowledgements, Software

MC2: Robin Donaldson – University of Glasgow www.brc.dcs.gla.ac.uk/software/mc2

Marcie: - BTU Cottbus <u>http://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Marcie</u>

Funding:

EU FP6 "Simulation Modelling of the Map Kinase Pathway" (SIMAP). EPSRC (UK), "Petri nets for multiscale Systems Biology"

### **Some references**

Qian Gao, Fei Liu, David Gilbert, Monika Heiner, David Gilbert (2011). A Multiscale Approach to Modelling Planar Cell Polarity in Drosophila Wing using Hierarchically Coloured Petri Nets, CMSB 2011

Robin Donaldson and David Gilbert (2008). A Model Checking Approach to the Parameter Estimation of Biochemical Pathways. CMSB 2008, LNCS 5307/2008, pp269-287.

Monika Heiner, David Gilbert, and Robin Donaldson (2008), **Petri Nets for Systems and Synthetic Biology**. Formal Methods for Systems Biology SFM 2008, LNCS 5016, pp. 215-264

David Gilbert, Monika Heiner and Sebastian Lehrack (2007). A Unifying Framework for Modelling and Analysing Biochemical Pathways Using Petri Nets CMSB 2007, LNCS/LNBI 4695, pp. 200-216.

David Gilbert and Monika Heiner, (2006). From Petri Nets to Differential Equations - an Integrative Approach for Biochemical Network Analysis, (ATPN06), LNCS 4024 / 2006, pp. 181-200

# **Demonstrations**

- Model: erk.param.spcontped
- Qualitative: Protein rises then falls P=?[(d[s\_4]>0)U(G(d[s\_4]<0))]
- Semi-qualitative: Protein rises then falls to less than 90% of peak concentration  $P=?[(d[s_4]>0)U(G(d[s_4]<0) \land F([s_4] < 0.9 * max[s_4]))]$
- Semi-quantitative:

Protein rises then falls to less than 90% of peak concentration by 60 minutes P=?[(d[s\_4]>0)U(G(d[s\_4]<0) ^ F(time = 60 ^ [s\_4] < 0.9 \* max[s\_4]))]

• Quantitative:

Protein rises then falls to less than  $1.6\mu$ Mol by 60 minutes P=?[(d[s\_4]>0)U(G(d[s\_4]<0) ^ F(time = 60 ^ [s\_4] < 1.6 ))]

David Gilbert