

# An Application of Model Checking to Epidemiology

Peter Drábik

Dipartimento di Informatica, Università di Pisa  
drabik@di.unipi.it

Guido Scatena

IMT Institute for Advanced Studies Lucca  
g.scatena@imtlucca.it



Institutions  
Markets  
Technologies  
**IMT**  
INSTITUTE  
FOR ADVANCED  
STUDIES  
LUCCA

## Goal

- To estimate suitability of dynamical sync-programs as a formalism to describe ecological and epidemiological models.
- To evaluate the applicability of probabilistic model checking as an analytic tool for understanding the dynamics of such models.

## Methods

### Modelling

Mathematical modelling of the progress of infectious diseases – to discover the likely outcomes of epidemics and to help manage them by vaccination.

- Standard method for large populations – deterministic approach using differential equations.

Individual-based methodology for small populations

- Pros: accounts for stochastic effects
- Cons: computationally expensive

Formalism – stochastic sync-programs

- Originally developed for description of biological systems such as signalling pathways [1].
- Each individual is modelled by a finite-state automaton.
- Interactions represented by using synchronisation referring to the moves of other automata.
- Multi-way synchronisation available.
- Sync-program is a parallel composition of sync-automata.
- Classical stochastic extension – interactions enriched with rates determining likelihood.
- Semantics in terms of continuous-time Markov chains (CTMC).

### Analysis

Probabilistic model checking

- Formal verification technique for analysis of systems that exhibit stochastic behaviour.
- Consists in establishing satisfaction of quantitative properties specified in probabilistic logic.
- Based on exhaustive exploration of the constructed model.
- Efficient tools available, here PRISM probabilistic model checker [2] is used.

Analyses performed

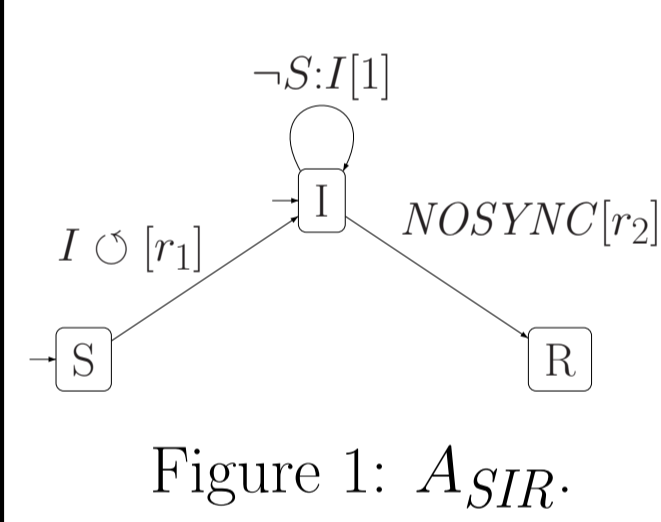
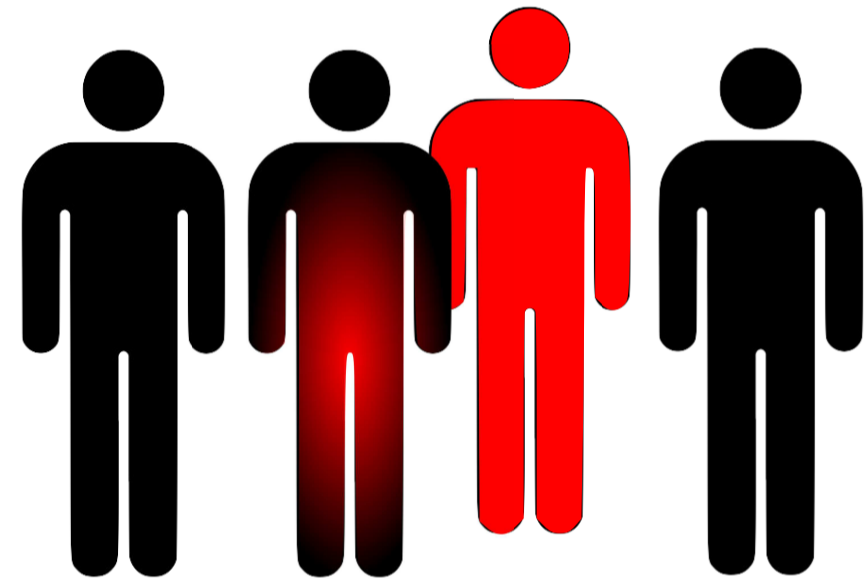
- Ad hoc translation of the stochastic sync-program to the PRISM input language is performed, preserving the CTMC semantics.
- Properties are specified in probabilistic logics PCTL and CSL.
- Properties are checked using the PRISM tool, by varying some parameters of the model.

### Case studies

Compartmental population models for many common childhood diseases that confer long-lasting immunity. Limitations – in order to obtain a finite-state model, we need to restrict the number of instances of automata of each type. In the application the limit is 10 for each type.

## SIR Model

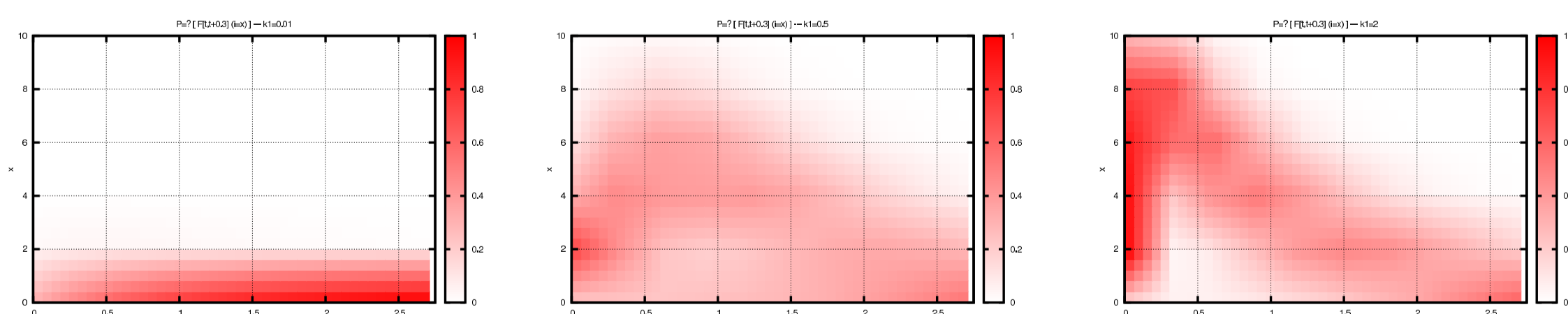
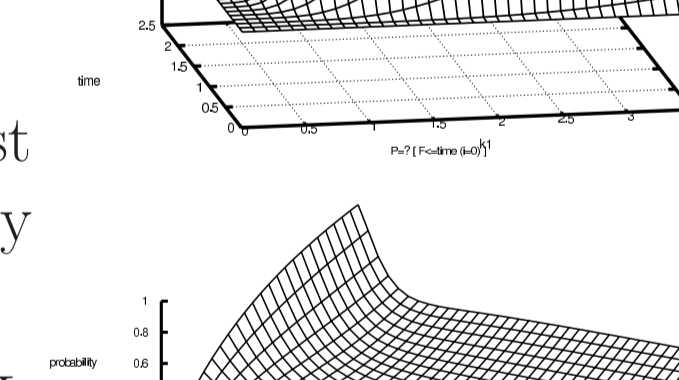
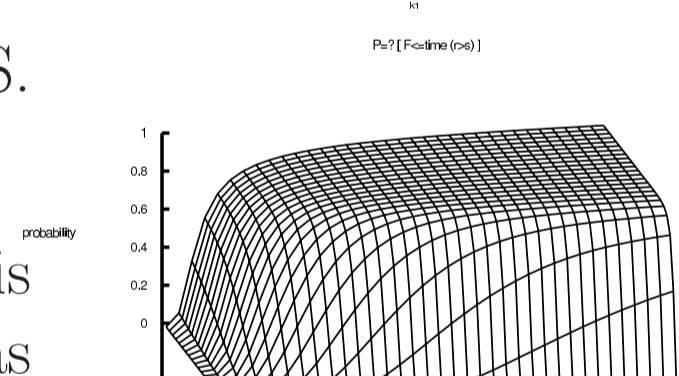
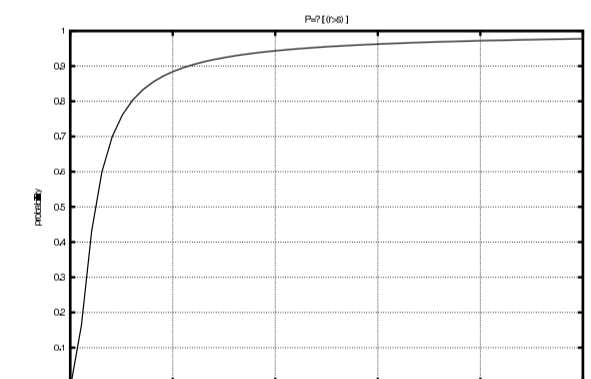
- Model for infectious diseases with droplet contact route of transmission such as measles, mumps and rubella.
- The population is divided into three compartments: those who are susceptible (S) to the disease, infected (I) and those who have recovered and are immune (R).
- Only hosts are modelled – each individual by one sync-automaton.



- The typical progress of each host is S to I to R.
- Atomic propositions:  $S$ ,  $I$  and  $R$ .
- States:  $\{S, \neg I, \neg R\}$ ,  $\{\neg S, I, \neg R\}$  and  $\{\neg S, \neg I, R\}$ .
- Move from S to I – getting an infection from an infected individual. Modelled by a synchronisation with a move of another sync-automaton that goes from state satisfying  $I$  to state satisfying  $I$  (denoted by  $I \odot$ ). Rate is  $r_1$ .
- Recovery from the disease occurs autonomously for each individual at rate  $r_2$ .

### Analysis results

- Program  $SIR = I || S || S || S || S || S || S || S || S || S$
- Different behaviour depending on the ratio  $R_0 = r_1/r_2$
- if  $R_0 > 1/S(0)$  – there is an epidemic outbreak with an increase of infectious population.
- if  $R_0 < 1$  – no outbreak occurs, independently of the initial population in S.
- Observed by fixing rate  $r_2$  and varying  $r_1$ .
- PCTL formula  $P = ?[(R > S)]$  – population of recovered individuals is bigger than population of susceptibles. At least half of the population was infected.
- First graph – if  $r_1$  is small, the probability of reaching a point, where at least half of the population was infected, is low. When increasing  $r_1$ , probability of such event increases towards 1.
- Second graph – this progress is plotted against time as the evaluation of CSL formula  $P = ?[F <= t((R > S))]$  again varying  $r_1$ .
- Probability of the retreat of the epidemic, or reaching state  $(I = 0)$  is 1. On the third graph it is shown how the retreat is likely to happen in time  $t$ .
- The bottom three figures express the probability of being in a state  $I = X$  at time  $Y$ , with  $r_1$  being 0.01, 0.5 and 2, respectively. First case – no outbreak, second and third cases – slow and rapid outbreak, respectively.



## SIR Model with Vectors

- Model for infectious diseases with vector-borne transmitted diseases.
- Demonstrates the dynamic aspect of the description language – creation of new automata in the runtime.
- Vectors are organisms that do not cause disease themselves but that transmit infection by conveying pathogens from one host to another.
- We believe these models can faithfully be employed for studying tick-borne encephalitis, Chikungunya (vector mosquitoes), Pappataci fever (vector sandfly) and diseases caused by Rickettsia bacteria like rickettsialpox, Boutonneuse fever and various spotted fevers (transmitted by ticks, fleas and lice).

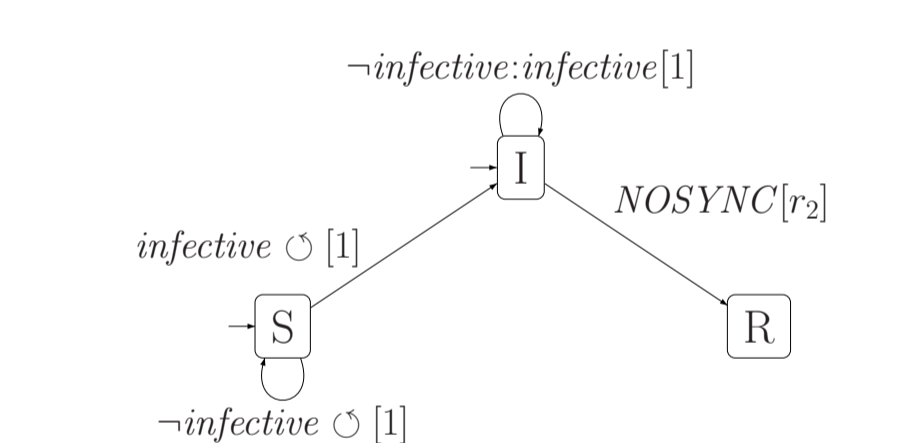
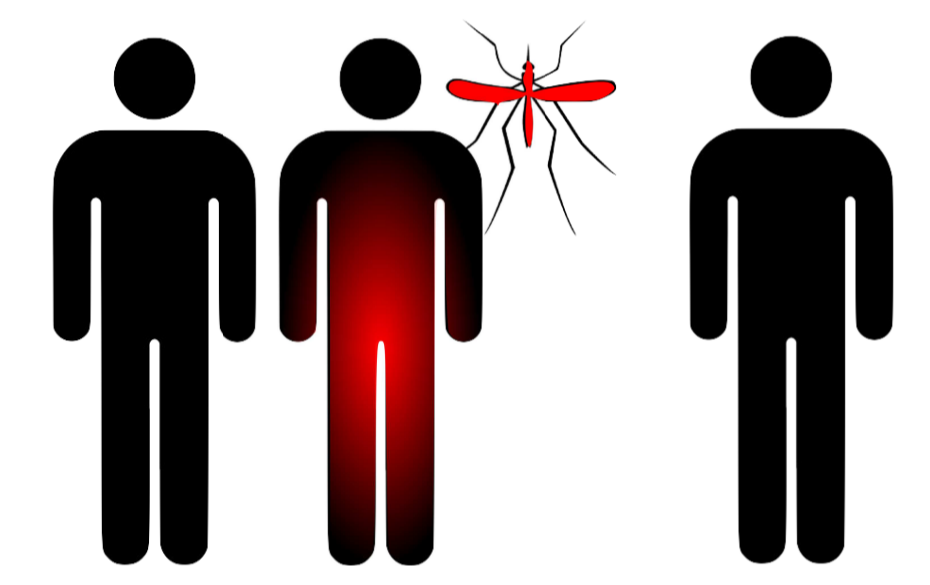


Figure 2:  $A_{host}$ .

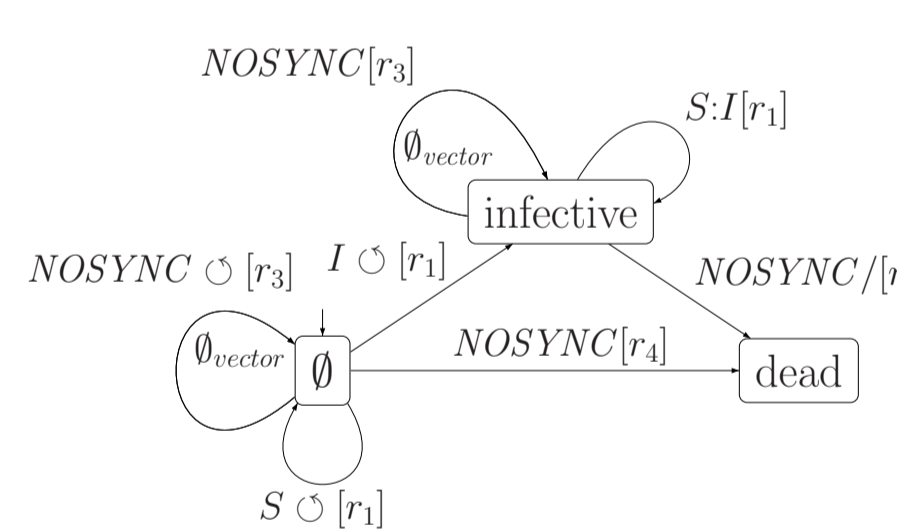
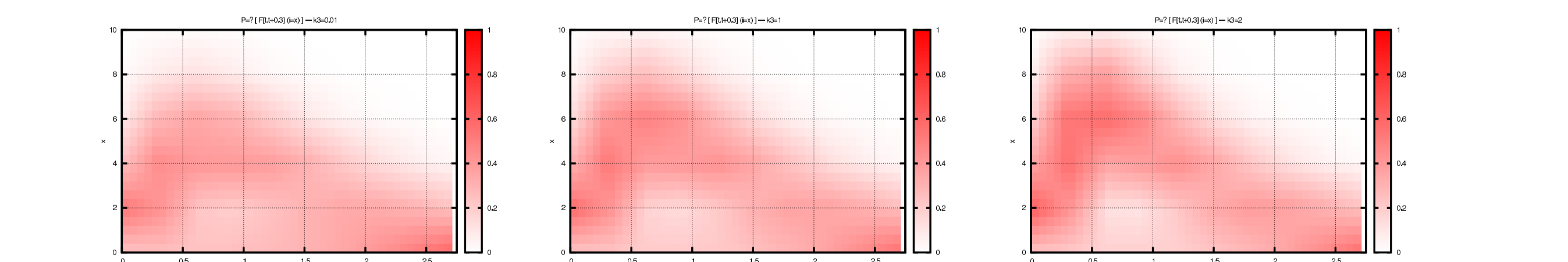


Figure 3:  $A_{vector}$ .

### Analysis results

- Program: nine susceptible hosts, one infected and five non infected vectors.
- Decisive role for the dynamics – creation rate of new vectors.
- Observed by varying  $r_1$  with  $r_2$  and  $r_4$  fixed.
- Retreat of the infection is unavoidable (probability of reaching  $(I = 0)$  is 1).
- First graph details its progress over time.
- Second and third graph are analogous to the ones in the model SIR.
- The probability distributions of values of I in time for  $r_3$  equal to 0.01, 1 and 2 are shown in the bottom figures.



## Conclusions

- We modelled a progress of an epidemic via a stochastic individual-based approach.
- All possible evolutions of the system considered providing exact results. In small populations the presence of stochastic effects can be witnessed.
- Stochastic sync-programs seem to be suitable means for description of these systems. Arbitrarily complex behaviour of one agent is expressible by a finite-state automaton. Interactions of individuals are specified by a general notion of synchronisation. Runtime automata creation allows for dynamicity.
- Probabilistic model checking was applied to study the properties of the model and sensitivity to parameters providing useful insight into the dynamics of the modelled system.
- Complex queries can be evaluated over the models by considering probabilities of values of variables in question and graphs can be plotted.
- In contrast to simulation approaches we obtain exact results based on inspection of all possible behaviours of the system, giving strong formal guarantees.
- A serious drawback are the computational costs of the procedure. In order to evaluate the queries in reasonable time (in the order of hours), we needed to limit the analysis to the order of tens of individuals.
- To combat this problem additional research is necessary.

## References

- [1] Peter Drábik, Andrea Maggiolo-Schettini & Paolo Milazzo (2010): *Dynamic Sync-programs for Modular Verification of Biological Systems*. Proc. of NCMA10. In press.
- [2] A. Hinton, M. Kwiatkowska, G. Norman & D. Parker (2006): *PRISM: A Tool for Automatic Verification of Probabilistic Systems*. In: H. Hermans & J. Palsberg, editors: Proc. of TACAS'06, LNCS 3920, Springer, pp. 441–444.

AMCA-POP 2010  
Jena, Germany  
25/8/2010