

# PEPA models of Epidemiology

Chris McCaig\*    Rachel Norman\*    Kevin O'Reilly\*  
Carron Shankland\*

## Extended Abstract

### Introduction

Increasingly, theoretical computer science techniques are being used to tackle biological problems. The question arises: which technique is best suited to the current problem? A related question is: which language features are required for describing biological problems? Our study specifically addresses the use of process algebra applied to epidemiology.

Having previously developed models of disease [1, 2] using WSCCS [3], we begin our investigations by considering the use of PEPA [4] for epidemiology. PEPA and WSCCS share the principle that systems are described in terms of *actions* and the *processes* that carry out those actions. Each has a relatively simple set of operations, including ordering of actions and choice between actions. Processes can be combined using parallelism, facilitating interaction.

WSCCS is a synchronous calculus, meaning that actions occur in every process at the same moment, in lockstep. The implication of this for modelling is that all processes must have some action available to them at each step, even if it is only the “tick” action, indicating a process which is idling. Time proceeds discretely, and there is no notion of the relation between each time step and real time. Choice between processes in WSCCS is influenced by weighted actions. For our epidemiological models it is convenient to represent the weights of choices as probabilities between 0 and 1. Communication between processes is essentially two-way, with a synchronisation occurring between an action and its co-action.

PEPA was chosen as a continuous time contrast to WSCCS. Actions occur at some *rate*, indicating length of time to perform the action, and yielding probabilistic choice between actions. Synchronisation in PEPA can be multi-way, although we have not yet exploited this in the models in this study. Comparisons are made with WSCCS on the process algebra, and via system dynamics expressed using differential equations. This allows further comparison with traditional models of disease [5]. Long term, other process algebras such as stochastic  $\pi$ -calculus [6] should be considered, as should other approaches to modelling, such as cellular automata and logic.

---

\*Department of Computing Science and Mathematics, University of Stirling, Stirling, FK9 4LA, UK. {cmc,ran,kor,ces}@cs.stir.ac.uk

```

S = (contact, c).S2;
S2 = (infect,p_i).I + (not_infect,1-p_i).S;
I = (transmit, t).I + (recover, p_r).R + (contact,c).I;
R = (lose_immunity, p_s).S + (contact,c).R;
Med = (transmit, infty).Med2;
Med2 = (contact, infty).Med + (decay, d).Med;
(S[900]<>I[100]) <transmit, contact> (Med[500])

```

Figure 1: Simple SIR model with indirect transmission in PEPA

## Basic SIR Model

The basic idea is to present an individual based model of disease spread. Observations about individual behaviour is used to abstract the general features of each class of individual and how they interact. There are three types of individual: S (susceptible), I (infected) and R (recovered). PEPA models naturally give rise to indirect transmission of disease, that is, where infection is passed to some sort of medium (the environment) and then a susceptible picks the infection up from the medium (rather than from the infected individual directly).

In the simple model presented in Figure 1 the behaviour is:

**S** Susceptible individuals may communicate with the environment, providing the possibility of being infected by the disease (S2). In the latter state, the individual may succumb to the disease and become I, or may avoid infection and remain S.

**I** Infected individuals pass their infection to the environment (transmit). In this model we take the view that the pool of infection is finite, and can be “used up” by being picked up by individuals who are not susceptible (as well as by those who are susceptible). Thus, Infected individuals may also pick up infection from the environment (contact), which has no effect on them. Infected individuals may recover from infection.

**R** Recovered individuals are immune from the disease for a time, returning to the susceptible pool. They may use up infection (in the same way as Infected individuals).

**Medium** The Medium is a pool of infection. It can accept infection (coming from Infected individuals). It can deliver infection to any of S, I and R. Infection also decays from the Medium.

An earlier version of this model had Med2 actively passing infection onto passive S, I and R, but this appeared not to give the usual infection dynamics (no epidemic resulted). This is worth further exploration, as is the translation between probabilities of the WSCCS model and rates of the PEPA model.

A broadly equivalent version of this model in WSCCS is presented in Figure 2. Note that in WSCCS the system is split into three distinct phases. In phase 1 contact is made, in phase 2 infection is transmitted from Infecteds to the Medium, and in phase 3 Infecteds may recover, and Susceptibles may become infected.

```

bs S1 1@1.contact:SI2 + 1.t:S2
bs I1 1@1.contact:I2 + 1.t:I2
bs R1 1@1.contact:R2 + 1.t:R2
bs Med_1 1.t:Med_2
bs Med2_1 1@1.contact^-1:Med_2 + 1.t:Med_2

bs S2 1.t:S3
bs SI2 1.t:SI3
bs I2 1.transmit^-1:I3 + 1.t:I3
bs R2 1.t:R3
bs Med_2 1.transmit:Med2_3 + 1.t:Med_3

bs S3 1.t:S1
bs I3 pr.t:R1 + (1-pr).t:I1
bs R3 ps.t:S1 + (1-ps).t:R1
bs Med2_3 1.t:Med2_1
bs Med_3 1.t:Med_1
bs SI3 pi.t:I1 + (1-pi).t:S1

basi L t
btr Population S1{s}|I1{i}|R1{r}|Med_1{e}/L

```

Figure 2: Indirect transmission in WSCCS.

## System Dynamics

The laws of process algebra specify the overall system behaviour in terms of transitions. Tools allow various forms of investigation including simulation, but this is often constrained by the ability to represent the state space. A more flexible approach is to generate (an approximation of the) population dynamics directly from the process algebra syntax. In WSCCS we can generate Mean Field Equations (Difference Equations). The WSCCS model of Figure 2 generates the following mean field equations:

$$\begin{aligned}
S_{t+1} &= S_t - \frac{p_i Med2_t S_t}{S_t + I_t + R_t} + p_s R_t, \\
I_{t+1} &= (1 - p_r) I_t + \frac{p_i Med2_t S_t}{S_t + I_t + R_t}, \\
R_{t+1} &= R_t + p_r I_t - p_s R_t. \\
Med_{t+1} &= Med_t + Med2_t - \frac{(Med_t + Med2_t) I_t}{Med_t + Med2_t + I_t} \\
Med2_{t+1} &= \frac{(Med_t + Med2_t) I_t}{Med_t + Med2_t + I_t}
\end{aligned}$$

which can be simplified since  $Med_t + Med2_t$  is constant.

In PEPA there are various ways to generate Ordinary Differential Equations [7, 8, 9], giving rather different results. For the model in Figure 1 the following ODEs are generated following the approach of [7]:

$$\begin{aligned}
dS/dt &= -c.\min(S, Med2) + (1 - p_i)S2 + p_s R , \\
dS2/dt &= c.\min(S, Med2) - S2 , \\
dI/dt &= p_i S2 - p_r I , \\
dR/dt &= p_r I - p_s R , \\
dMed/dt &= c.\min(S, Med2) - t.\min(I, Med) + d.Med2 , \\
dMed2/dt &= -c.\min(S, Med2) + t.\min(I, Med) - d.Med2 ,
\end{aligned}$$

The other methods provide slightly different terms for interacting components, removing the *min* function and replacing it with the number of active components only, with a side condition guaranteeing there are passive components with which to interact.

## Further Work

A more thorough comparison of the capabilities of PEPA for modelling epidemiology remains to be carried out. Topics to consider include:

**Direct Transmission** Biologically, several different forms of transmission have been identified: directly transmitted, e.g. measles and chickenpox; sexually transmitted, e.g. HIV; vector borne (mosquitos, ticks, etc.), e.g. malaria, Lyme disease, encephalitis; transmission through free living infective stages, e.g. bovine TB; transmission on a network i.e. via local and long range interactions; transmission via superspreaders, e.g. cholera, SARS. Is it necessary to represent all of these differently? Can all diseases be considered as indirect transmission since there must be some medium present (air, bodily fluids, etc)?

Our WSCCS models express both direct and indirect transmission of disease [2], and the other categories above are deemed to fit these two types of transmission.

**Births and Deaths** In a closed population disease is fairly straight-forward. More interesting factors can be considered once births and deaths are added to the population.

**Frequency Dependent and Density Dependent Transmission** An advantage of an individual based model is that it is possible to vary the way an individual interacts with others in the population, and to investigate the impact this has on transmission rates. Consider an individual interacting in their population. A disease modelled using frequency dependent transmission indicates that the individual is likely to maintain the number of contacts they have, no matter the size of the population. This is common with e.g. sexually transmitted diseases. Conversely, a disease represented using density dependent transmission indicates that the individual makes contacts proportional to the population, therefore, in a more dense population, the number of contacts is higher, and the likelihood of passing on the disease becomes greater.

Our WSCCS models have been able to express both frequency and density dependent transmission of disease [2], although frequency dependent transmission is what arises most naturally from WSCCS models.

**Control** From a biological point of view, once the dynamics of a disease is understood the next question is how to limit the effects of that disease through e.g. vaccination, quarantine, culling. (Alternatively, if disease is being used as a control itself, e.g. of crop pests, the question is how to maximise the effect of the disease, while minimising environmental impact). The model presented here had individuals recover from the disease becoming immune to infection for a time, and then returning to the pool of susceptible individuals. Questions such as vaccination and quarantine remain to be addressed.

At Stirling we are about to begin a three year project *System Dynamics from Individual Interactions: A process algebra approach to epidemiology* supported by the EPSRC which will look at these questions and others.

### Acknowledgement

CM is supported by an EPSRC Doctoral Training Account. KO was supported by a Nuffield Foundation Undergraduate Research Bursary, and a Carnegie Trust Studentship.

## References

- [1] R. Norman and C. Shankland, “Developing the use of process algebra in the derivation and analysis of mathematical models of infectious disease,” in *Computer Aided Systems Theory - EUROCAST 2003*, vol. 2809 of *Lecture Notes in Computer Science*, pp. 404–414, Springer-Verlag, 2003.
- [2] C. McCaig, *Disease and WSCCS*. PhD thesis, 2007. Unpublished Draft.
- [3] C. Tofts, “Processes with probabilities, priority and time,” *Formal Aspects of Computing*, vol. 6, pp. 536–564, 1994.
- [4] J. Hillston, *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.
- [5] R. Anderson and R. May, “The population-dynamics of micro-parasites and their invertebrate hosts,” *Philosophical transactions of the Royal Society of London Series B*, vol. 291, pp. 451–524, 1981.
- [6] C. Priami, “Stochastic  $\pi$ -calculus,” *The Computer Journal*, vol. 38, pp. 578–589, 1995.
- [7] J. Hillston, “Fluid Flow Approximation of PEPA models,” in *QEST’05, Proceedings of the 2nd International Conference on Quantitative Evaluation of Systems*, pp. 33–42, Torino: IEEE Computer Society Press, September 2005.
- [8] J. Hillston and J. Bradley, “Quantitative analysis of pepa models through continuous state space approximation.” Unpublished Draft, 2007.
- [9] Various, “Pepa eclipse plugin.” Current implementation of ODE analysis, 2007.