

# Algebra and logic for practical systems modelling: an overview

David Pym

HP Labs, Bristol and

University of Bath

[david.pym@hp.com](mailto:david.pym@hp.com), [d.j.pym@bath.ac.uk](mailto:d.j.pym@bath.ac.uk)

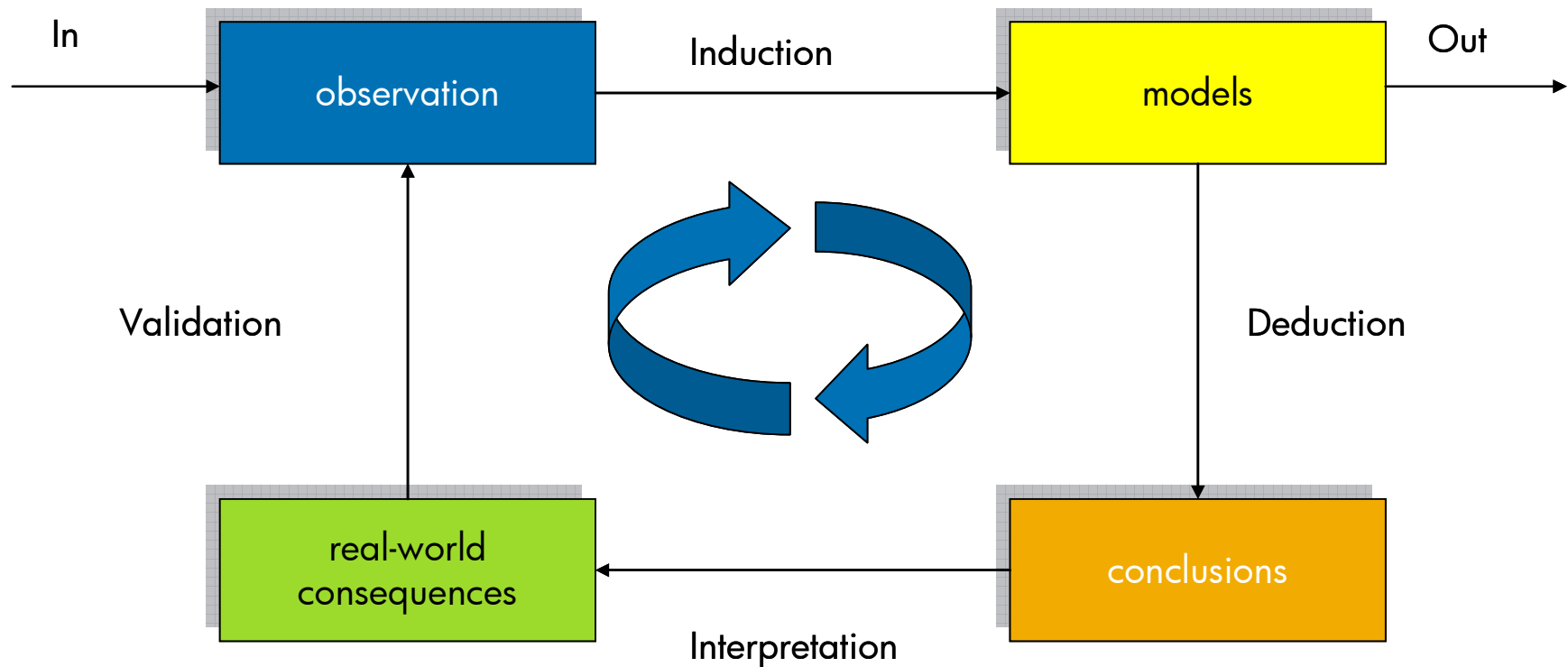
# A brief introduction to HP Labs

- HP's corporate advanced research facility, distinct from R&D in the businesses.
- Long-term, disruptive brief, but clearly connected to the business.
- About half in Palo Alto, CA, about a quarter in Bristol, UK, and other labs in Tokyo, Haifa, Bangalore, St. Petersburg, and Princeton (part of Bristol).

# A modelling philosophy

- Consider techniques from mathematical computing-related sciences: process algebra, logic, probability theory, queuing theory.
- Develop and deploy these ideas as modelling tools in the sense of classical applied mathematics.
- *Some* heritage from Demos2k.

# The classical modelling cycle



# The components of systems models

- Systems are distributed, so **location** and **connectivity** are important; think network.
- **Resources** are distributed around the network.
- **Processes** execute relative to resources and location.
- Systems exist inside **environments**, from which events are incident.

# Aside: connections to economics

- Why? Because systems are deployed to deliver services satisfying economic incentives.
- What? Performance objectives, security objectives, cost objectives.
- How? One connection derives from the stochastic aspects of the modelling. More later.

# Our approach

- Model each of these components — location ( $L$ ), resource ( $R$ ), process ( $E$ ), environment — essentially independently, and combine in process algebra,

$$L, R, E \rightarrow^a L', R', E',$$

and in logic,

$$L, R, E \vDash \varphi$$

- Environment handled stochastically on top of the algebra (and maybe the logic).

# Processes and Resources

- Starting point for processes is Milner's SCCS (omitting recursion for now):

$$E ::= a \mid a:E \mid E + E \mid E \times E \mid \nu R.E$$

- Here  $\nu R.E$  replaces restriction, with resources  $R$  forming a (possibly) ordered monoid:

$$R = (R, \circ, e, \leq)$$

- For example, non-negative integers with  $+$ ,  $0$ , and less-than-or-equals.

# Operational form for 'SCRIP'

- So what we have here is a situation in which resources and processes co-evolve:

$$R, E \rightarrow^a R', E'$$

where  $R' = \mu(a, R)$ , where  $\mu: \text{Act} \times R \rightarrow R$ , with coherence conditions.

- The basic action rule is

$$R, a:E \rightarrow^a \mu(a, R), E$$

# Some inference rules

- Product:

$$\frac{R, E \xrightarrow{a} R', E' \quad S, F \xrightarrow{b} S', F'}{R \circ S, E \times F \xrightarrow{ab} R' \circ S', E' \times F'}$$

- Hiding:

$$\frac{R \circ S, E \xrightarrow{a} R' \circ S', E'}{R, \nu S.E \xrightarrow{\nu S.a} R', \nu S'.E'}$$

- Sum: as you'd expect.

# Bisimulation

- Bisimulation is more awkward than one at first imagines (and we got it wrong at first).
- There are two versions:
  - Local,  $R, E \approx R, F$ , an equivalence relation, not a congruence (took us a while to notice), but many useful meta-theoretic properties all the same.
  - Global,  $E \sim F$ , a congruence, some useful theory, but a significant weakness.

# Tools

- Demos2k (Birtwistle, Christodolou, Tofts): captures process and environment, limited resources, no location.
- Located Demos2k (Monahan, Collinson, Pym): adds location to the structural part of Demos2k.
- Gnosis 'knowsys' (Monahan, Collinson, Pym): addresses the full location, resource, process, and environment requirement.

# A modal logic, 'MBI'

- Builds on the bunched logic BI, which is also the basis for separation logic.
- BI freely combines additives (cf. classical and intuitionistic logic) and multiplicatives (cf. linear logic) to give a logic with an entertaining 'resource semantics'.
- Resource semantics is based on (ordered) monoids, just like SCRP.
- Semantically,  $\circ$  captures combination of resources, with unit  $e$ , and  $\leq$  captures comparison of resources.

# Basic resource semantics

- Additive conjunction,  $\wedge$ :

$$R \vDash \varphi \wedge \psi \text{ iff } R \vDash \varphi \text{ and } R \vDash \psi$$

- Multiplicative conjunction,  $*$ :

$$R \vDash \varphi * \psi \text{ iff there are } S \text{ and } T \text{ such that}$$

$$S \circ T \leq R, \text{ and}$$

$$S \vDash \varphi \text{ and } T \vDash \psi$$

- Similar situation for implication (and units).

# Hennessy-Milner-style resource semantics

- Basic judgement form is

$$R, E \vDash \varphi$$

- So the action clause is

$R, E \vDash \langle a \rangle \varphi$  iff there is a transition  
 $R, E \xrightarrow{a} R', E'$  such that  $R', E' \vDash \varphi$

- And so the product clause goes like

$R, E \vDash \varphi * \psi$  iff there are  $S$  and  $T$  such that  
 $S \circ T \leq R$  and  $F$  and  $G$  such that  $E \sim F \times G$ ,  
and  $S, F \vDash \varphi$  and  $T, G \vDash \psi$

- Note alternative choices for  $\leq$  and  $\sim$ .

# Other Logical Stuff

- Other connectives include:
  - Multiplicative implication;
  - Multiplicative modalities;
  - Additive and multiplicative quantifiers.
- Can be formulated intuitionistically or classically.

# Basic meta-theory

- The Hennessy-Milner equivalence theorem relates logical equivalence and bisimulation:
  - $E \approx F$  iff  $E \not\equiv \Vdash F$  holds for a suitable fragment of MBI (multiplicative implication and multiplicative modalities are problematic);
  - $E \sim F$  implies  $E \not\equiv \Vdash F$  holds for all of MBI; but
  - $E \not\equiv \Vdash F$  implies  $E \sim F$  is an interesting question:
    - As stated, cannot hold, since  $\not\equiv \Vdash$  is not a congruence, but maybe for some relation ‘between’  $\approx$  and  $\sim$ ?
    - Problematic in our setting, because of the delicacy of the interaction between equivalence the substructural connectives;
    - It seems that  $\langle a \rangle$  makes too few distinctions.
- Generally, the difficulty is if/where to put the universal quantification over resources.

# Adding location

- We model location just as we model resource, but identifying a suitable mathematical structure.
- For location, we identify the following:
  - A set of **places**;
  - Directed **connections** between places;
  - A notion of **sub-place**;
  - Connectivity-respecting **substitution**.
- Directed graphs provide a leading example.
- We can then reconstruct SCRIP and MBI to make use of location. Actions are located, processes aren't:  $\mu: \text{Act} \times \mathbf{R} \times \mathbf{L} \rightarrow \mathbf{R} \times \mathbf{L}$ .

# Applying the Calculus and Logic to Systems Security

- In fact, suppress location for now.
- We can add an operator for roles:  $E \blacktriangleright F$ .
- This is a (non-commutative) concurrent composition:

$$\frac{R, F \xrightarrow{a} R', F' \quad R \circ S, E \xrightarrow{a} R' \circ S', E'}{R \circ S, E \blacktriangleright F \xrightarrow{a} R' \circ S', E' \blacktriangleright F'}$$

where  $F$  is simulated by  $E$ .

- Corresponds to a logical modality, ‘says’,  $\{E\}\varphi$ :
- $R, G \vDash \{E\}\varphi$  iff there is an  $F$  such that
  - $R, G \approx (R, E \blacktriangleright F)$ ,  $E$  simulates  $F$
  - and  $R, F \vDash \varphi$ .

# Systems Security Examples

A range of security examples, including:

- Chains of trust, guards;
- Co-signing;
- Mutual exclusion;
- Roles.

These examples tend to require a fair amount of space to write out the details. I can supply them offline, or see papers at <http://www.cs.bath.ac.uk/~pym/recent.html>.

- But also some 'large scale' examples.

# The tools and examples

- Demos2k:
  - The ‘boats’ model;
  - Problems with generalizing the ‘boats’ model;
  - A generic issue: location.
- Gnosis (Gk. *knowledge*, think ‘know sys’):
  - Rigorous re-implementation;
  - Adds *location* as a first-class construct.

# Demos2k: 'boats'

```
cons arrival=negexp(10.0);
cons dock=20.0;
cons unload=normal(14,3);
cons tug=3; Cons jetty=2;
cons simdur=24*60;

class boat= {entity(boat,boat,arrival);
  getR(jetties,1);
    getR(tugs,2);
    hold(dock);
    putR(tugs,2);
      hold(unload);
  getR(tugs,1);
    hold(2.0);
    putR(tugs,1);
    putR(jetties,1);
} (**boat**)

entity(boat,boat,0.0);
  hold(simdur);
  close;
```

# An end-to-end applied study: USB memory stick security

- Modelling the Human and Technological Costs and Benefits of USB Memory Stick Security. Proc. WEIS 2008, Dartmouth College, NH: <http://weis2008.econinfosec.org/papers/Pym.pdf>
- End-to end study:
  - Empirical study of memory stick usage and security policy compliance by Merrill Lynch employees; leading to
  - Process model (using Demos2k, to be replaced by Gnosis) of memory stick lifecycle; providing
  - Evidence of confidentiality–availability trade-off; explained by
  - A Central-Bank-style (macro-)economic model (impulse-response under utility maximization).
- A more theoretical economics paper in FC&DS 09.

# Ongoing work and directions

- The Gnosis tool: further developments (of location implementation, of an experiment manager, etc.).
- A systems modelling book based on our existing work + the tool, all with a security-driven flavour.
- Game-theoretic views of information flow/access control in systems models.
- Model-checking tools.
- Understanding how to integrate systems models and economic models: economic models, such as those in information security economics, make use of 'system parameters'.

# Acknowledgements

- Matthew Collinson, Chris Tofts, Brian Monahan, Mike Yearworth, and, at least historically, Peter O'Hearn.
- Adam Beutement, Robert Coles, Jonathan Griffin, Christos Ioannidis, Angela Sasse, and Mike Wonham.
- Martin Sadler.
- The Royal Society of London.
- The University of Bath.
- The UK Technology Strategy Board.