Facilitating the Comparison of Social Simulations using E-Science

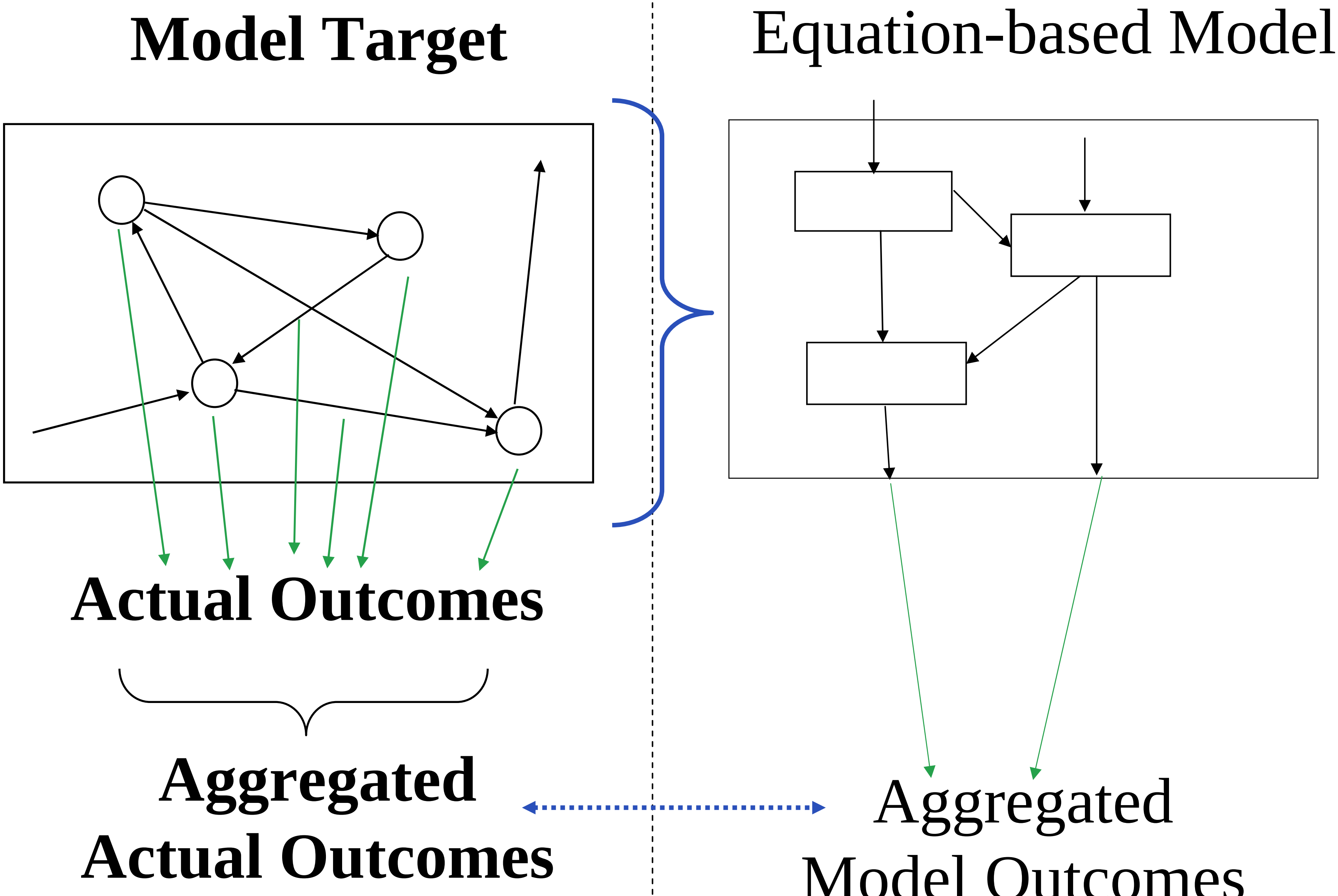
Bruce Edmonds

Centre for Policy Modelling

Manchester Metropolitan University

Part 1: Introduction *– about social simulation*

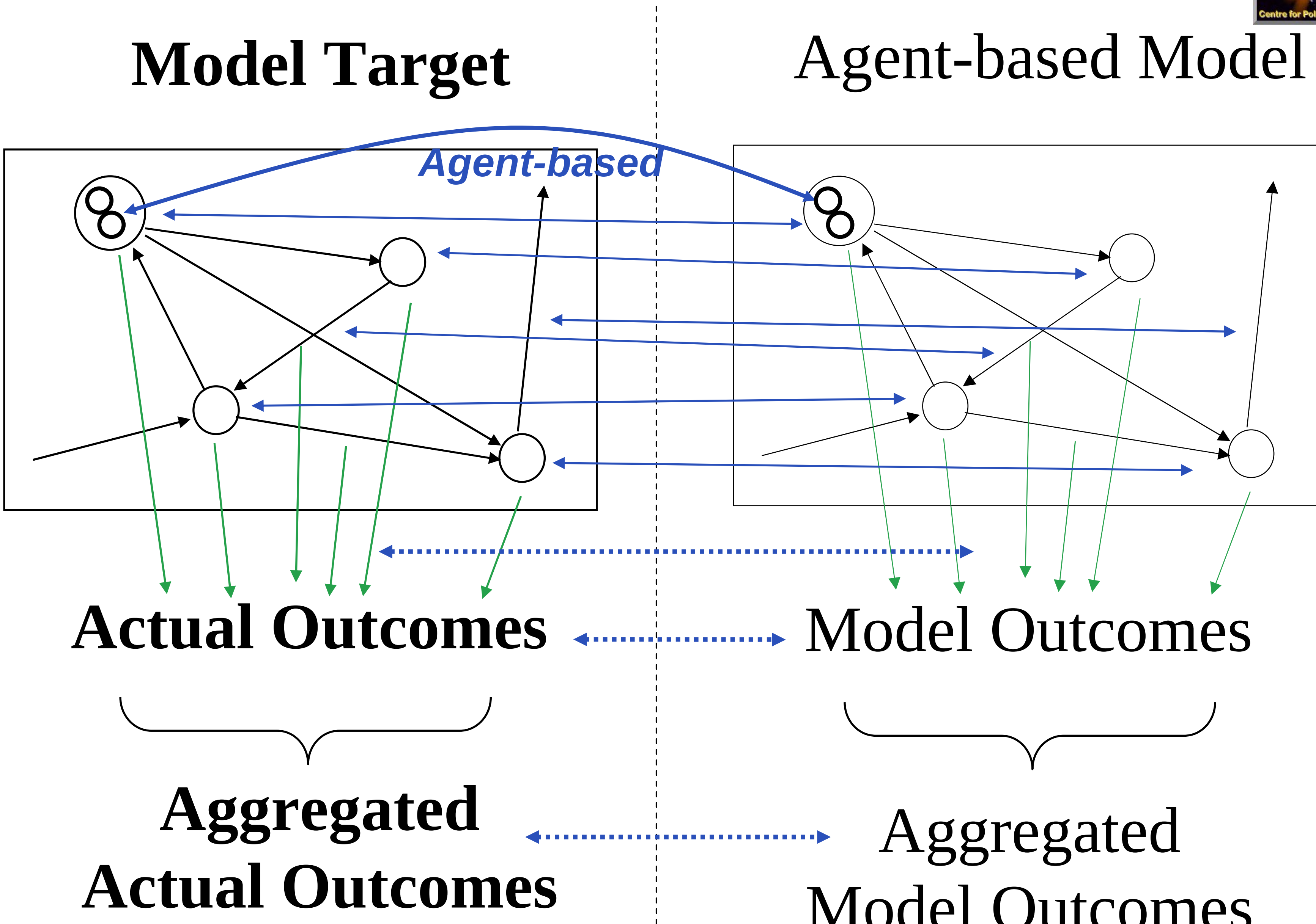
Equation-based modelling



Properties of (equation-based) Mathematical Models

- Long tradition, many techniques/results
- Central use of numbers and proof
- In very simple cases can derive closed-form (i.e. general) conclusions…
- …but more often, in practice, they are calculated ("numerically simulated")
- Essentially about states
  - Atemporal (where time occurs it is reified)
  - Inference can works in may ways
- Assumptions necessary to represent world
- 'Art' of approximation and application

Individual/agent-based simulation



Properties of (agent/individual-based) Simulations

- Short tradition, fewer techniques/results
- Central use of algorithms and computation
- Difficult to derive general conclusions
  - More like an experiment than an inference
- Essentially about process
  - Temporal directionality
  - Process of unfolding observable
- More representational in practice
  - In time and in composition
- More suggestive of interpretation

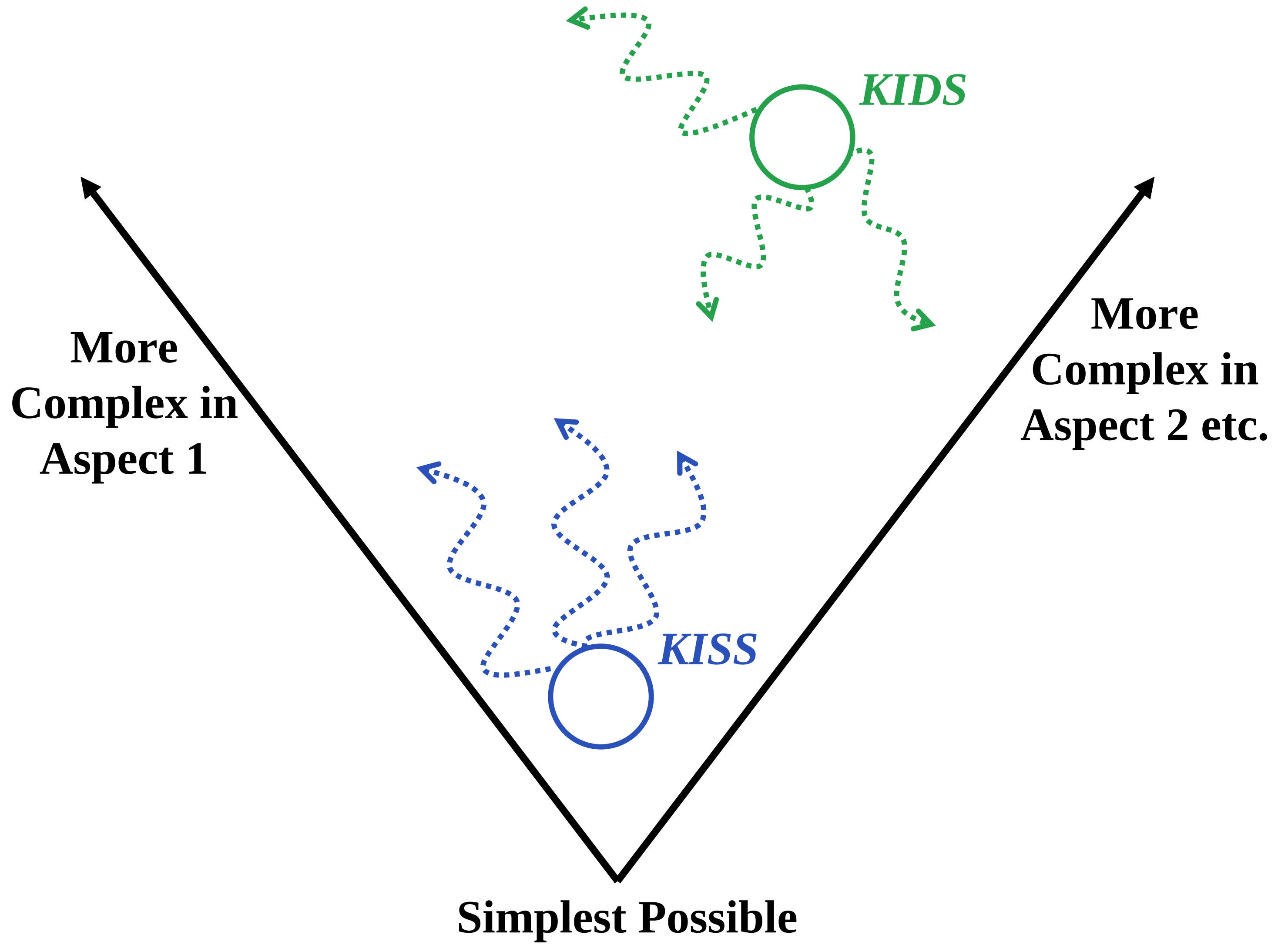
KISS – "Keep it simple stupid!"

- Simulation is as simple "as possible"
- Only elaborate the simulation in steps where simulation is found to be inadequate
- Simulations tend to be very abstract and "physics like"
- Uses strong assumptions/simplifications
- Aims for generality
- Weakly related to phenomena via conceptual model of simulator (analogy)

KIDS – "Keep it descriptive stupid!"

- Simulation is as descriptive as information, techniques and resources allow for
- Then adapted (simplified or changed) as becomes necessary
- Simulations are intricate and context-specific (does not aim for generality)
- Related to phenomena in as many ways as possible (including to anecdotal evidence)
- More amenable to detailed criticism and comparison by stakeholders and experts

KISS vs. *KIDS* illustrated



Part 2: Comparing Simulations

Why Compare Simulations?

KISS approach

- Generality of approach to be tested
- Simulation tested against its interpretation

KIDS approach

- Experiments to test understanding of simulation
- Discovering when it best matches the observed

All

- Finding bugs and unintended behaviours
- Find out which of the different aspects are responsible for differences

Some aspects of a simulation

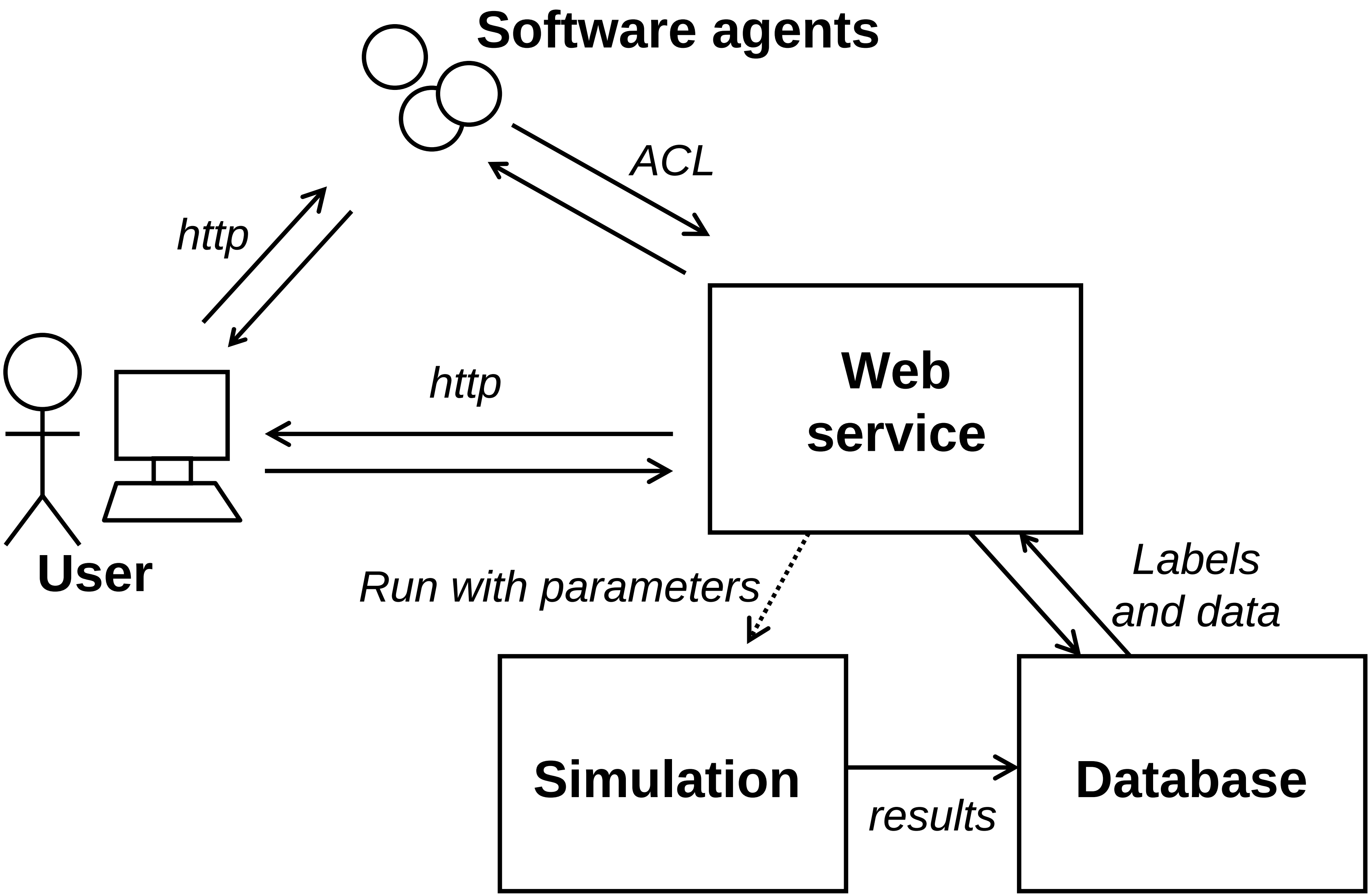
- The *core* algorithm
- Some other known aspects of the algorithm included simply to facilitate the simulation
- Some other aspects of the algorithm that the original simulator did not know existed
- The particular parameter values chosen for the simulation runs
- The computational environment in which the code is run
- The particular way in which the behaviour of the simulation was analysed or summarised
- Our understanding (or theory) about the simulation's workings

Possible Approaches

1. Disseminating precise descriptions of simulations to allow their replication
2. Disseminating results arising from runs of the code with given parameters
3. Disseminating the code plus instructions
4. Allowing others to run some reference code remotely and query the results
5. Allowing others to automatically find and interact with some reference code remotely

Part 3: A pilot – *a quick outline and some thoughts*

Structure



Software agents
ACL
http
http
User
Web service
Run with parameters
Labels and data
Simulation
results
Database

Summary of pilot structure

- Online querying of database
- Database contains: simulation names, parameters, parameter ranges, run labels, result sets, state/stage of runs
- Offline queue of simulation runs requested by users produces results sets cached in database
- Web service interfaces with users (and *theoretically* software agents)
- No web semantic for simulation internals

(Very limited) lessons learnt

- Very quickly unmanageably long queues of simulation requests build up, thus management of computational resources is very important, e.g. by (notional) charging
- In the short (and maybe medium) term automatic use via semantic web/agents by users will be difficult and limited
- Simple but predictable and reliable is more important than theoretical power for users

Part 4: Discussion *– the symbol grounding problem and the GRID*

The “Symbol Grounding Problem”

- Steven Harnad (1990)
- That however intricately symbols are related to each other they can not acquire meaning (wrt to the world) except by:
  1. Human interpretation filling the gap
  2. Learning by interaction with the world
- Given lack of (2) it will be very difficult for the semantic web to provide meaning to labels separate from human interpretation

The case with simulations

- The meaning of parts of a simulation is:
  - Crucial to the use of a simulation
  - Often contested by simulators
  - Often derives heavily from context
  - Thus terms have a (limited but vague) domain of legitimacy
- Thus the outlook for the semantic side of the semantic web/GRID is not good
- Rather tools for human manipulation of systems and tokens is more promising

The End

Bruce Edmonds

bruce.edmonds.name

Centre for Policy Modelling

cfpm.org