Validation and simulation: engineering simulations for science

Fiona A. C. Polack

Department of Computer Science, University of York, UK

EPSRC grants EP/E053505/1 and EP/E049419/1
www.cosmos-research.org

EmergeNET April 2010

Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
  - Modelling at individual level
  - Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
  - Throw-away simulations that only address one question
  - Naive simulations that scientists cannot use effectively
  - Simulations that are not validated w.r.t. science

Focus on validation

Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
  - Modelling at individual level
  - Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
  - Throw-away simulations that only address one question
  - Naive simulations that scientists cannot use effectively
  - Simulations that are not validated w.r.t. science

Focus on validation

Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
  - Modelling at individual level
  - Inherent modelling in time and space
- Problem: "opaque thought experiments" [Bullock et al]
  - Throw-away simulations that only address one question
  - Naive simulations that scientists cannot use effectively
  - Simulations that are not validated w.r.t. science

Focus on validation

Complex systems simulation for science

Simulation: A model executed on a computer to reveal time-dependent behaviour

- e.g. Agent-based models to support scientific research
  - Modelling at individual level
  - Inherent modelling in time and space
- Problem: “opaque thought experiments” [Bullock et al]
  - Throw-away simulations that only address one question
  - Naive simulations that scientists cannot use effectively
  - Simulations that are not validated w.r.t. science

Focus on validation

Validating simulations

- Validation seeks to show that the right system has been built
- Validity is determined with respect to that purpose [Sargent 2005]
- Validation guidance exists for conventional simulations
  - but does not adapt directly to complex systems

Conventional simulation validation 1

Animation, Face validity, Turing test	Evaluation of observables: outputs and parameters
Comparison, historical or predictive validation	Against *valid* analytical models or other simulations
Degenerate and extreme tests, parameter variability, sensitivity analysis	Domain-style testing under normal and extreme conditions
Event validity	Compare events in real and simulated systems
Internal validity	Consistency of results across runs

[after Sargent 2005]

Conventional simulation validation 2

Combination of "traditional methods":	Validate assumptions, compare to sound theory, and do empirical validity checks

[after Sargent 2005]

Traditional approaches:

- Rationalism: Assumptions are rationally justifiable; valid models arise from valid assumptions
- Empiricism: Assumptions and outcomes are empirically validated
- Positive economics: The model can predict the future, so causal relationships and mechanisms are of no concern

Complex simulation validation: summary of problems

- No existing validated models to compare with
- Existing understanding of event, behaviours is limited
- Very different systems give similar observables
  - ... and the same systems has different observables in different situations

So, use multiple validation approaches
Validation becomes relative not absolute

Complex simulation validation: summary of problems

- ◦ No existing validated models to compare with
- ◦ Existing understanding of event, behaviours is limited
- ◦ Very different systems give similar observables
  - ... and the same systems has different observables in different situations

So, use multiple validation approaches

Validation becomes relative not absolute

Complex simulation validation: summary of problems

- No existing validated models to compare with
- Existing understanding of event, behaviours is limited
- Very different systems give similar observables
  - ... and the same systems has different observables in different situations

So, use multiple validation approaches
Validation becomes relative not absolute

Arguing validity

- Establish the basis for believing that a simulation is an adequate model
  - relative to its purpose, criticality and impact
- Summarise evidence to support the argument
- Expose the argument to critical review

Techniques from critical systems engineering can help:

- Deviational analysis, flaw hypothesis
- Argumentation development and representation

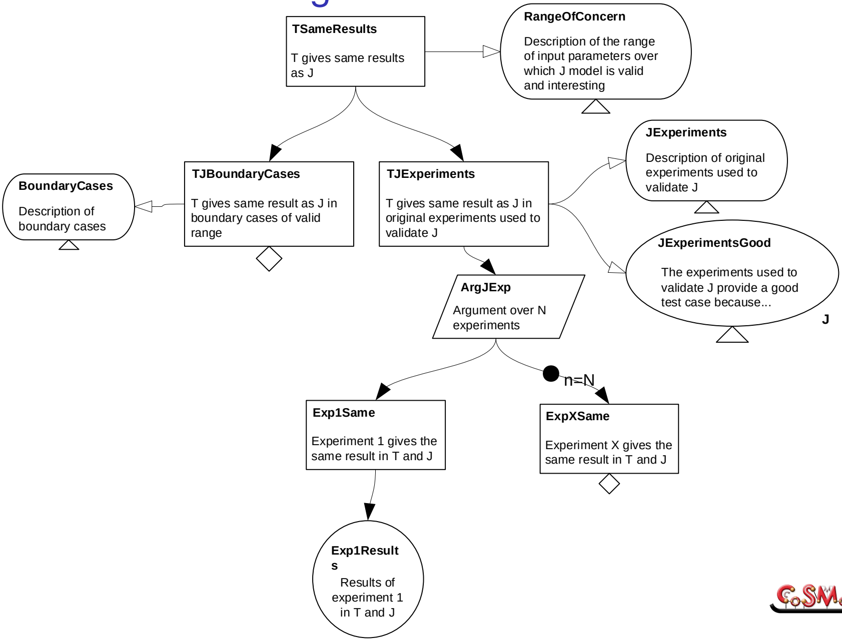
Arguing validity

- Establish the basis for believing that a simulation is an adequate model
  - relative to its purpose, criticality and impact
- Summarise evidence to support the argument
- Expose the argument to critical review

Techniques from critical systems engineering can help:

- Deviational analysis, flaw hypothesis
- Argumentation development and representation

Part of a validation argument



Divide and conquer: three aspects of validity

- *Engineering validation*: conventional development and verification activities on the simulation product
  - exposes design decisions, implementation compromises
- *Calibration*: tuning parameters, behaviours, scales of operation
  - exposes assumptions, abstractions, simplifications and mappings
- *Scientific validation*: demonstrate the scientific use of the simulation, for instance by simulating real experiments and comparing results

Argument of validity over these

- Sufficient to convince those who need to be convinced!

Divide and conquer: three aspects of validity

- *Engineering validation*: conventional development and verification activities on the simulation product
  - exposes design decisions, implementation compromises
- *Calibration*: tuning parameters, behaviours, scales of operation
  - exposes assumptions, abstractions, simplifications and mappings
- *Scientific validation*: demonstrate the scientific use of the simulation, for instance by simulating real experiments and comparing results

Argument of validity over these

- Sufficient to convince those who need to be convinced!

Purpose revisited: criticality and impact

Careful definition of purpose (domain, domain model)

- Scientific purpose identifies domain of interest
- Engineering purpose expressed in a domain model
- domain and domain model determine criticality and impact of simulation

Criticality and impact

- *Criticality* relates to the role of the simulation in context.
  - A speculative exploration of possible factors has low criticality
  - A simulation that aims to contribute scientific understanding has high criticality.
- *Impact* is relative importance of simulation findings to research area
  - where few research tools exist or the area is new, impact is usually high
  - if simulation findings are subsequently backed by traditional research, impact is low

If impact or criticality is high, then explicit validation of the engineering and the science must be planned

Criticality and impact

- *Criticality* relates to the role of the simulation in context.
  - A speculative exploration of possible factors has low criticality
  - A simulation that aims to contribute scientific understanding has high criticality.
- *Impact* is relative importance of simulation findings to research area
  - where few research tools exist or the area is new, impact is usually high
  - if simulation findings are subsequently backed by traditional research, impact is low

If impact or criticality is high, then explicit validation of the engineering and the science must be planned

Criticality and impact

- *Criticality* relates to the role of the simulation in context.
  - A speculative exploration of possible factors has low criticality
  - A simulation that aims to contribute scientific understanding has high criticality.
- *Impact* is relative importance of simulation findings to research area
  - where few research tools exist or the area is new, impact is usually high
  - if simulation findings are subsequently backed by traditional research, impact is low

If impact or criticality is high, then explicit validation of the engineering and the science must be planned

Summary

- Complex systems simulation need to be validated, but traditional methods are not directly applicable
- Instead, construct an argument over evidence of why the simulation is believed to be valid for the purpose
- Focus on issues relating to engineering, calibration and sensitivity, as well as assumptions and design decisions
- Do what is merited by the purpose, criticality and impact of the simulation
  - Aim to convince those using the simulation and its results
  - And expose the basis for belief in validity to review and critique