Noname manuscript No.
(will be inserted by the editor)

The CoSMoS Process

Simulations for Scientific Exploration

Paul S. Andrews · Susan Stepney · Jon Timmis · Fiona Polack · Adam Sampson · Peter Welch · Frederick Barnes

Received: date / Accepted: date

Abstract The use of computer simulations to investigate the behaviour of complex systems is an increasingly common discipline. We present here the outline of a process (the CoSMoS process) we are developing to aid and guide the construction of such simulations.

Keywords Complex systems · agent-based simulation · process

Computer simulation is fast becoming an invaluable tool in many areas of scientific enquiry. We are interested in how best to construct such simulations that aid theory exploration, hypothesis generation and suggest real-world experimentation. Our approach focuses on agent-based simulations in which the agents of a system are directly modelled as computational processes, allowing complex behaviours to emerge naturally from their simulated interactions. We are developing a generic process for engineering agent-based simulations that is being extracted from various simulation case-studies developed with scientists from a broad range of disciplines (immunology, ecology, sociology). In addition we are extracting patterns that describe how and when to achieve each step of the process. We are developing a massively concurrent simulation infrastructure, and investigating how to argue validity of simulations for scientific rigour. We emphasise that the CoSMoS process is an *agile* approach, adaptable both to a variety of problems and to changing circumstances during simulator construction and use.

The CoSMoS process identifies a number of named *roles*: the *modeller*, the *implementor*, the *experimenter* and the *domain expert*, who are involved in various distinct *phases* of the process. These phases are the *discovery* phase, the *scoping* phase, the *implementation* phase and the *experimentation* phase.

Discovery Phase – An understanding of the *domain of interest* (the real-world system under study) and the goal of the simulation is established. This phase requires the interaction between the modeller and the domain expert (someone who has specific domain knowledge not possessed by the other roles). The initial task is to establish the

This work is part of the CoSMoS project, funded by EPSRC grants EP/E053505/1 and EP/E049419/1. www.cosmos-research.org

P. Andrews
Department of Computer Science, University of York, UK

goal of the simulation exercise: what aspects of the domain are you going to simulate and why. This identifies the level of abstraction and accuracy to which you are required to model the domain. An explicit *domain model* is then constructed by the modeller to capture the understanding of the domain of interest at the identified level of abstraction, identifying domain components, component behaviours, and any uncertainties in our knowledge of the domain. The domain model is documented using any suitable language or approach, such as diagrammatic representations of behaviours and tables of abstractions and assumptions. Importantly the domain model will include any observable emergent or expected behaviours of the system, including their relationship to the system components/agents.

Scoping Phase – From a domain model and goal, the simulator is scoped to establish the foundation for its construction, use and evaluation. This phase requires the input from all the named process roles. First, the key aspects of the discovery phase are distilled: a description of any emergent/expected behaviours; guidelines for measuring these behaviours; any hypotheses, questions or predictions to be investigated using the simulator. Next, the technology to be used to construct the simulator is identified (the *infrastructure domain*, along with the effects this choice has on the simulator architecture. Other elements to scoping include: any limitations of the work e.g. issues of scale and minimum or realistic numbers of agents; the need and level of any *validity argumentation*; and lastly *success criteria* to measure whether goals have been achieved.

Implementation Phase – Once scoped, a simulator is designed and implemented by the implementor and modeller. The design is captured in a *platform model* (again documented using any suitable language or approach), the creation of which is driven by the concepts identified in the domain model and the chosen infrastructure domain. The key tasks involved in creating the platform model include: removal of emergent behaviours (so that they may emerge in simulation); mapping of domain model components to agents and environment; a design implementation for agents and environmental factors; addition of instrumentation to measure emergent behaviours, visualise and interact with simulations. The simulator results from the direct implementation (coding) of the platform model.

Experimentation Phase – The first job of the experimenter is to *calibrate* the simulator by adjusting parameters to match our knowledge of the real world system. This provides a baseline for simulations and affords a better understand of the relationship between the real world and what the simulations show. Next, the experimenter performs explorative and/or predictive *in silico* experimentation. The processes of calibration and further experimentation proceed along similar lines. First, predictions for specific simulation behaviours are produced from the details identified in the scoping phase, and simulations are instantiated to run experiments that test such predictions. During calibration, the outcomes feed back into the simulator or highlight deficiencies in the earlier modelling stages. During experimentation, the simulation outputs are used to construct an *analysis model*. In the same way that the domain model is an explicit representation of our knowledge of the domain, the analysis model is an explicit representation of the behaviour of the simlator. The analysis model may include a statistical model generated from simulation results and/or a description of the visual artefacts observed in the simulator. The analysis model then forms the basis for comparison by the experimenter and domain expert to the success criteria identified during the scoping phase.