

ManyCell: A multi-scale simulator for cellular systems

Joseph O. Dada¹, and Pedro Mendes^{2,3}

Short Abstract — Biological systems are composed by a hierarchy of scales. A common representation takes cells as the basic units which are organized in larger structures: cultures, tissues and organs. Within cells there is also a great deal of organization, both structural (organelles) and biochemical (pathways). To better understand these systems there is a need for computational environments that allow for a mapping of functional behaviors across the scales. Multiscale simulations of tissues or cell cultures should reveal macroscopic properties that are a consequence of the internal organization in their component cells. Here we present ManyCell, a multiscale cell simulation software environment.

Keywords — simulation, multiscale, modelling, software, ODE, agent-based.

I. INTRODUCTION

ONE of the main aims of multiscale simulation in biology is to be able to simulate entire organs or even entire organisms at various levels of detail. Different methods have been implemented, some taking a mean-field approach, others representing the entire hierarchy of entities [1]. We feel that while the objective is to represent entire organs, this should be able to simulate all of its cells and their intracellular pathways. However this means that one may need to include $10^6 - 10^{12}$ cells. The software ManyCell, that is described here, is intended to achieve this requirement and implements technical solutions that contribute to numerical efficiency and scalability. At the same time the software uses the most natural representations for these phenomena: ordinary differential equations (ODEs) are used to model the internal biochemistry of cells; each cell is modelled as an agent, where its state changes depend on discrete events that are triggered by the ODEs. The system also allows other entities, such as the extracellular medium or other extrinsic factors, to be modelled as agents.

II. IMPLEMENTATION

To be able to deal with very large numbers of cells, the software must be a) scalable, meaning that it is not limited by the finite resources of a single computer, and b) time efficient. ManyCell includes technologies to specifically

address these issues. To improve efficiency we have adopted an In Situ Adaptive Tabulation (ISAT) mechanism [2]. To provide scalability the system uses distributed computing such that arbitrary numbers of processing units can be added to share the work. Unlike other solutions that use object-oriented programming to represent agents, we adopted a new solution where a relational database management system (RDBMS) is used to manage the agents. This allows the number of transactions (event triggers) to be extremely large, and as a bonus the simulation results are easily managed and queried. Agents themselves require simulating ODE-based pathway models, which is carried out by the software COPASI [3], each instance being able to be run on a different processing unit. Because of this, ManyCell readily accepts models encoded in SBML [4], such as those in the BioModels database [5]. The tissue/culture organization rules are encoded in a simple XML format. Communication between agents is carried out using Web Services.

III. RESULTS AND DISCUSSION

We present simulation experiments of proliferating yeast cell cultures where the base model for each cell is from Chen *et al.* [6], modified to consume a medium substrate for its growth. We tested three cases a) serial – where one cell at a time solves its ODEs, b) parallel – where two or more cells solve their ODEs simultaneously; iii) ISAT – where the ODE solutions make use of the ISAT scheme. We observe that the ISAT scheme provides some 50-fold acceleration. Using only two computers (one for ODE calculations, the other for the RDBMS) we were able to simulate growth up to 10^6 cells. Thus simulations can easily be scaled up to run at least 10^9 cells in computer farms or cloud computing. The ManyCell architecture allows scaling up the number of computers dedicated to numerics, as well as those dedicated to the data management (RDBMS). This system is thus expected to be able to simulate entire organs by representing all of its cells.

REFERENCES

- [1] Dada JO, Mendes P (2011) *Integrative Biology* **3**, 86-96.
- [2] Pope S (1997) *Combustion Theory Modelling* **1**, 41-63.
- [3] Hoops S, et al. (2006) *Bioinformatics* **22**, 3067-74.
- [4] Hucka M, et al. (2003) *Bioinformatics* **19**, 524-531.
- [5] Le Novère N, et al. (2006) *Nucleic Acids Res.* **34**, D689-91.
- [6] Chen KC, et al. (2004) *Mol. Biol. Cell* **15**, 3841-62.

Acknowledgements: This work was funded by EU FP7 grant 201142, UniCellSys.

¹School of Computer Science, The University of Manchester, UK. E-mail: joseph.dada@manchester.ac.uk

²School of Computer Science, The University of Manchester, UK. E-mail: pedro.mendes@manchester.ac.uk

³Virginia Bioinformatics Institute, Virginia Tech, USA. E-mail: mendes@vt.edu