OpenVT: an introduction and discussion

Randy Heiland

PhysiCell Project

Nov 10, 2023

OpenVT – A Standardized Ecosystem for Virtual Tissue Simulation

2-year NSF grant (ends 2025-8-31)

Specific aims include: development of shared standards for specification of VT models between CompuCell3D and PhysiCell, creation of cell-type description libraries, standardization of initial conditions, and description of model outputs. These standards will then be used to define the needed APIs that allow the interconnection and reuse of models.

https://new.nsf.gov/funding/opportunities/pathways-enable-open-source-ecosystems-pose

PIs: Glazier, Wild, Macklin   https://www.nsf.gov/awardsearch/showAward?AWD_ID=2303695

OpenVT will enable sharable, cross-platform modeling tools and shareable model specifications. OpenVT aims to accelerate the understanding of complex biological mechanisms related to tissue development, homeostasis, and disease. The project will focus on agent-based modeling (ABM) approaches, where tissues and organs are constructed using discrete cells, coupled with subcellular network models of signaling, gene regulation and metabolism and partial differential equations that simulate the extracellular movement of oxygen, growth substrates, signaling factors, and therapeutic compounds. Specific aims (previous slide) while also providing a concrete technical roadmap to support the integration of additional open source ABM VT frameworks in the future. This approach allows for the integration of subcellular, cell-level, and tissue-level phenomena, providing explanatory power and enabling high-precision virtual experiments. The project will support the integration of other VT frameworks and the creation of educational and distribution facilities to enable their widespread adoption and extension of OpenVT. The OpenVT ecosystem aligns with successful community-driven initiatives in scientific software development, such as the Systems Biology Markup Language (SBML) project. It aims to foster collaboration, establish standards for model inputs and outputs, and provide user support and training. By transitioning to an open-source ecosystem, the project aims to reduce duplication of effort, promote software and model sharing, and democratize access to modeling capabilities. Project outcomes will be available at OpenVT.org

CompuCell3D [1] and *PhysiCell* [2]—developed by two co-PIs of this proposal Dr. Glazier and Dr. Macklin, respectively—are at the forefront of the multi-cell cell-based tissue modeling community. Both platforms have a wide user base and complement each other by implementing different underlying tissue modeling methodologies.

Other flagship packages in the domain include *Morpheus* [3], *Biocellion* (closed-source) [4], *TissueForge* [5], *CHASTE* [6], *BioDynaMo* [7], and *Simmune* [8].

Specific Aims

1. standards for specification of VT models between CompuCell3D and PhysiCell
2. creation of cell-type description libraries
3. standardization of initial conditions
4. description of model outputs
5. define APIs to allow interconnection and reuse of models
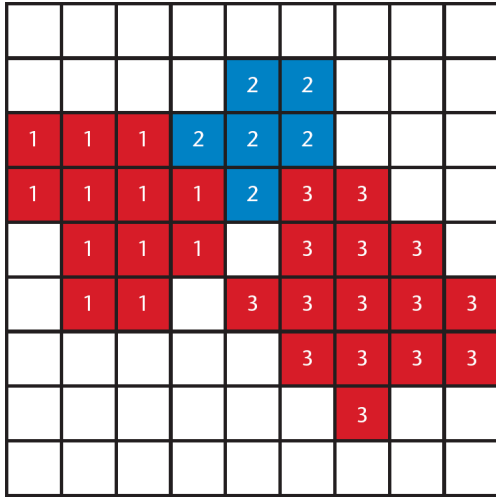
Scope of the work: ABMs

Cellular (agents, PDEs) + Subcellular ([SBML wiki](#))
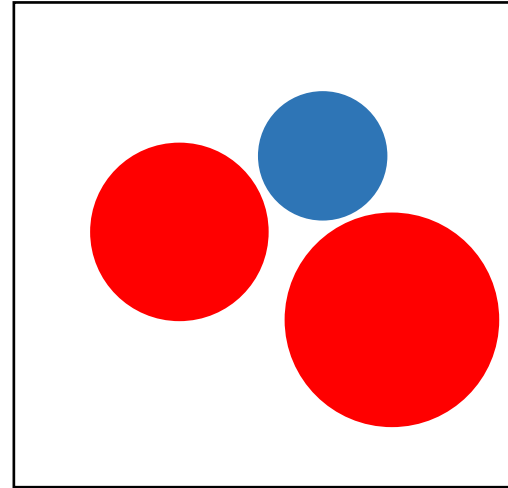
Let's prioritize the tasks for OpenVT:

- Cellular: agents
    - How are they defined? (for each ABM framework)
    - Data format to represent both? ("standard")

Agents: spatial definition

- CC3D: contiguous pixels on a lattice
- PhysiCell: center + volume (lattice-free)

ECM = 0

Specifying agents initial conditions (geometry)

- ## CC3D: PIFF (Potts Initial File)
  https://compucell3dreferencemanual.readthedocs.io/en/latest/pif_initializer.html

  cell# celltype x1 x2 y1 y2 z1 z2

- ## PhysiCell: CSV

  https://github.com/MathCancer/PhysiCell/blob/master/changes.md#1110

  x,y,z,celltype[,…]

Standard data format for cell geometry?

What makes a "standard"?

(open discussion: 2-3 mins)

Questions (1)

Assume we create a new, standard data format.

- What about backwards compatibility for each framework?

- Will each framework continue to support both their original format and the new one? And eventually phase out the original?

- How do we reach a consensus on the format?

- XML-based? (and XML Schema?) JSON5?

Questions (2)

- How do we encourage adoption of a standard?

- Should we provide separate tool(s) to create ICs? (If so, should be cross-platform)

- What's the general structure of the format? Propose:

  - XML defining the multicellular objects/params; ref. data files (rf. MultiCellDS slide)

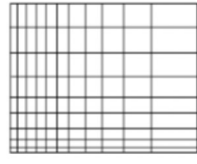- Concerns mapping weird CPM cell morphology to center-based?

Some possible data formats

- VTK - vtk.org

  - examples.vtk.org/site/VTKFileFormats/

  - examples.vtk.org/site/VTKBook/05Chapter5/

- Blender - blender.org, wiki.blender.org/wiki/Main_Page

  - https://unstructured-grids.readthedocs.io/en/latest/ug.html

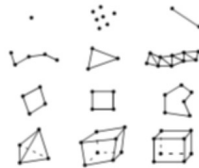# VTK data formats
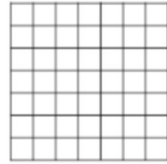
(a) Image Data
(b) Rectilinear Grid
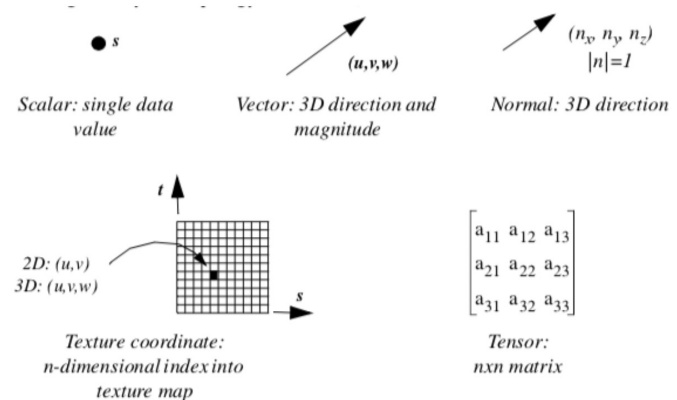(c) Structured Grid
(d) Unstructured Points
(e) Polygonal Data
(f) Unstructured Grid

5.5 Attribute Data

Attribute data is information associated with the structure of the dataset. This structure includes both the dataset geometry and topology. Most often, attribute data is associated with dataset points or cells, but sometimes attribute data may be assigned to cell components such as edges or faces. Attribute data may also be assigned across the entire dataset, or across a group of cells or points. We refer to this information as attribute data because it is an attribute to the structure of the dataset. Typical examples include temperature or velocity at a point, mass of a cell, or heat flux into and out of a cell face.

Scalar: single data value

Vector: 3D direction and magnitude

Normal: 3D direction

2D: (u,v)
3D: (u,v,w)

Texture coordinate: n-dimensional index into texture map
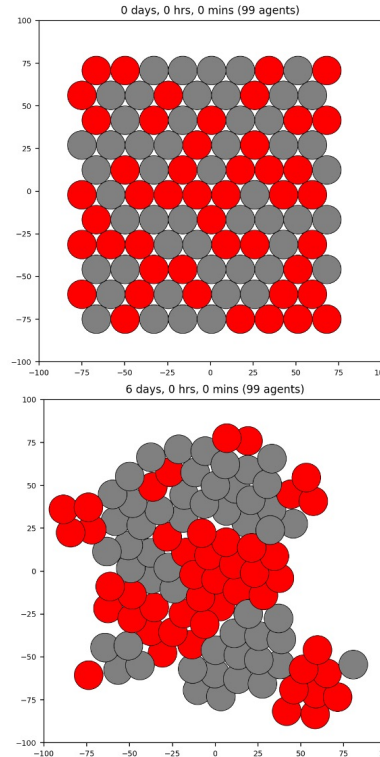
Tensor: nxn matrix
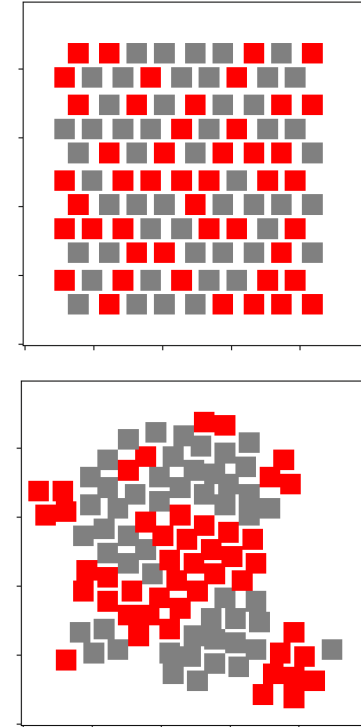
Very preliminary work

Python script to create a VTK (.vtu) file from PhysiCell cells ICs.
Python script to create native format for CC3D.

Converting a later snapshot as ICs.

PhysiCell

CC3D

- https://github.com/OpenVT/playground/issues/2

MultiCellDS (MultiCellular Data Standard)

- A proposed standard for [ABM] multicellular data

- http://multicellds.org/ (Note: not https)

- XML format

- "A novel part of MultiCellDS is the digital cell line"

- Currently being used by PhysiCell (but references data files in .mat (MATLAB v4) format)

OpenVT community

- Should an online community exist on its own?
  If so, how?
  - Slack? GitHub? Google-group?

  https://github.com/OpenVT (GitHub organization)
  https://groups.google.com/g/morpheus-users, good example

- OpenVT.org – must exist (in the grant)