What it is
What it could be
What happened to Biology at the end of XX\textsuperscript{th} century?

A NEW APPROACH TO DECODING LIFE:
Systems Biology

Trey Ideker\textsuperscript{1,2}, Timothy Galitski\textsuperscript{1}, and Leroy Hood\textsuperscript{1,2,3,4,5}

Institute for Systems Biology\textsuperscript{1}, Seattle, Washington 98105; Departments of

Ohmsha, Ltd. and Springer-Verlag

invited paper

Perspectives on Systems Biology

Hiroaki KITANO
Sony Computer Science Laboratories, Inc.

©Ohmsha, Ltd.
What happened to biology at the end of XX\textsuperscript{th} century?
New way of doing biomedical research

Needs for interplay between models and reality tests

Needs for systems thinking and integration of heterogeneous knowledge

Needs for cooperation and standardisation

\[ \frac{dx}{dt} = f(X, P, t) \]
Computational modelling left the niches


- **Pharmacometrics models** Labrijn et al. Therapeutic IgG4 antibodies engage in Fab-arm exchange with endogenous human IgG4 in vivo. *Nat Biotechnol* 2009

- **Physiological models** Noble. Modeling the heart from genes to cells to the whole organ. *Science* 2002; Izhikevich and Edelman. Large-scale model of mammalian thalamocortical systems. *PNAS* 2008

Computational models on the rise

BioModels Database growth since its creation

http://www.ebi.ac.uk/biomodels/
Interest from new stakeholders

- “Biologists”: computational models look “useful”, “serious”

- Publishers: computational are respectable, and can be published in high profile journals

- Funding agencies: Models could help with the major challenges (read “science that can be sold to citizen/electors”): Health, Food, Energy...

- Industries: Models could help with the major challenges (read “new opportunities to make money”): Pharmas, crops, biofuels ...
The matrix of standards for M&S in Sys Bio

<table>
<thead>
<tr>
<th></th>
<th>Model descriptions</th>
<th>Simulations and analysis</th>
<th>Numerical results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal requirements</td>
<td><img src="miriam.png" alt="MIRIAM" /></td>
<td><img src="miase.png" alt="MIASE" /></td>
<td></td>
</tr>
<tr>
<td>Data-models</td>
<td><img src="sbml.png" alt="SBML" /></td>
<td><img src="sedml.png" alt="SEDML" /></td>
<td><strong>NuML?</strong></td>
</tr>
<tr>
<td>Terminologies</td>
<td><img src="sbgn.png" alt="SBGN" /></td>
<td><img src="kisao.png" alt="KISAO" /></td>
<td></td>
</tr>
</tbody>
</table>
Dimension 3: Covering alternative modelling approaches
Parallel and redundant efforts

- Neurobiology
- Physiology
- Systems Biology
- Pharmacometrics
- Developmental biology, plant biology
- BioPAX
- SbML
- SbGN
- NineML
- FieldML
- IncF
- Virtual Physiological Human network of excellence
- Drug Disease Model Resources
What if the world-wide web was built like this?
The correct way to do it

Personal info: vCard
Presentation: CSS
Semantics: RDF
Graphics: SVG
Display: HTML
Covering the entire modeling in the life-sciences

model type

Mean-field

state transition

process description

model semantics

biological semantics

visual representation

level of representation

model life-cycle

model generation

model description

simulation analysis

numerical results

...
Existing standards interoperability
Threats to the whole enterprise

- Current efforts are largely dependent on key people. Their disengagement means stalling or disaggregation.
- Current funding structure is fragile. Many different grants, sometimes only supporting meetings, none of them infrastructure rolling funding, often tied to individuals.
- Current efforts, being developed under the umbrellas of specific institutions are not immune against intellectual property claims that would harm the community.
- Existing standards are developed with very different approaches, quality checks, and are based on completely different assumptions (e.g. implicit knowledge versus explicit mathematics).
- APIs, converters etc. need industry-grade support, incompatible with standard academic usages and possibilities.
Overarching standardisation structure

The “WorldWide Web consortium” of modelling in biology

http://co.mbine.org/
Mission 1: Coordinating the standards

- **CORE STANDARDS:** Efforts fulfilling COMBINE criteria and aiming at following COMBINE rules and interoperate with other COMBINE standards.

- **ASSOCIATED EFFORTS:** Standards that are not representation formats, but aiming at enrich or bridge the core standards.

- **RELATED EFFORTS:** Formats developed by other communities, that complement or interoperate with COMBINE formats, and that we would like to see joining COMBINE or collaborating closely to COMBINE.
Current core COMBINE standards

- Model semantics, Model structure, Process description: S8ML

- Models semantics, Simulation and Analysis: SEDML

- Biological semantics, Model structure, Process description, Entity relationships: BioPAX

- Visual representation, Model structure, Process description, Entity Relationships, Activity flow: GN8
Current associated standardization efforts

- Concept and data reference:
  - MIRIAM Registry
  - Identifiers.org URIs

- Terminologies:
  - Modeling
  - Simulation
Related standardization efforts

PSI-MI

FieldML

NuML?
COMBINE does NOT aim to take over the development of the standard formats, but help coordinating and supporting this process.
Mission 2: Coordinating meetings

- Annual COMBINE meetings
  - COMBINE 2010: October 6–9, Edinburgh, 81 attendees
  - COMBINE 2011: September 3-7, Heidelberg, 82 registrations
  - COMBINE 2012: End of summer, beginning autumn, Toronto
  - COMBINE 2013: Date unknown, location unknown
    http://www.surveymonkey.com/s/combine-harmony-hosting-interest

- The Hackathons on Resources for Modeling in Biology
  - HARMONY 2011: April 18-22, New-York City, 59 attendees
  - HARMONY 2012: Date unknown, location unknown
    http://www.surveymonkey.com/s/combine-harmony-hosting-interest
Mission 3: Developing Standard Operating Procedures

- Technical requirements
  - Who format covers what portion of the modeling space
  - Which technical solutions exist and must be used by the formats
  - How do formats interface
  - How to specify and document formats
  - ...

- Governance
  - How to initiate and maintain standardisation efforts
  - How to communicate with users and developers
  - How to develop a democratic and robust governance
  - ...
First tentative SOP: guidelines to develop a core COMBINE standard

- List the new development as a related standardization effort
- Join COMBINE community and attend meetings
- Comply with COMBINE criteria
  - Must cover aspects of modeling significantly different from the existing set of COMBINE standards
  - Must be described in precise technical specification documents and formal specification languages
  - Specifications and other materials must be publicly available free of charge to everyone and be unencumbered by licensing restrictions
  - Development must be open. The entire COMBINE community must be able to participate without exclusion
  - Must be developed and used by more than a single team or organisation.
  - Development process must be led by democratically elected editorial boards
  - Mature software support must exist, including standard API implementations, and possibly validation tools
  - Development must be stable and active
- Decision by the coordinators (alt: vote of the community?)
Mission 4: Recognised voice

- COMBINE aims to become a “standardisation” body
  - This means a quality label. A “COMBINE standard” is a guarantee of stability, community endorsement, support etc.
  - COMBINE production can be used in SOPs at other organisations
  - COMBINE must be an actor on par with FGED, PSI, INCF etc.
- Single point of contact with user organisations including Industry
  - Tool developers (General platforms or specific tools)
  - Publishers
  - Pharmaceutical industry
- A point of contact for funding bodies
- A point of contact for legal entities, e.g. government and regulatory bodies
Where to find more information?

Communities
- http://biopax.org/
- http://sbgn.org/
- http://sbml.org/
- http://sed-ml.org/
- http://biomodels.net/
- http://biomodels.net/kisao
- http://biomodels.net/sbo
- http://biomodels.net/teddy
- http://biomodels.net/miase
- http://biomodels.net/miriam

Coordination
- http://combine.org/

Acknowledgements

SBML editors: Frank Bergmann, Andrew Finney, Stefan Hoops, Michael Hucka, Nicolas Le Novère, Sarah Keating, Chris Myers, Sven Sahle, Herbert Sauro, Jim Schaff, Lucian Smith, Darren Wilkinson

SBGN editors: Emek Demir, Nicolas Le Novère, Huaiyu Mi, Stuart Moodie, Falk Schreiber, Anatoly Sorokin, Alice Villéger

BioPAX editors: Peter D'Eustachio, Oliver Ruebenacker, Andrea Splendiani

SED-ML editors: Richard Adams, Franck Bergmann, Nicolas Le Novère, Andrew Miller, David Nickerson, Dagmar Waltemath

Metadata: Mélanie Courtot, Nick Juty, Camille Laibe, Anna Zhukova

The whole community of Computational Systems Biology
<table>
<thead>
<tr>
<th></th>
<th>Model descriptions</th>
<th>Simulations and analysis</th>
<th>Numerical results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Minimal requirements</strong></td>
<td><img src="#" alt="MIRIAM" /></td>
<td><img src="#" alt="MIASE" /></td>
<td></td>
</tr>
<tr>
<td><strong>Data-models</strong></td>
<td><img src="#" alt="SBML" /></td>
<td><img src="#" alt="SEDM" /></td>
<td>NuML?</td>
</tr>
<tr>
<td></td>
<td><img src="#" alt="GN" /></td>
<td><img src="#" alt="BioPAX" /></td>
<td></td>
</tr>
<tr>
<td><strong>Terminologies</strong></td>
<td><img src="#" alt="SSO" /></td>
<td><img src="#" alt="KISAO" /></td>
<td><img src="#" alt="TEDDY" /></td>
</tr>
<tr>
<td></td>
<td>Model descriptions</td>
<td>Simulations and analysis</td>
<td>Numerical results</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Minimal requirements</strong></td>
<td><img src="image" alt="MIRIAM" /></td>
<td><img src="image" alt="MIASE" /></td>
<td></td>
</tr>
<tr>
<td><strong>Data-models</strong></td>
<td><img src="image" alt="SBML" /></td>
<td><img src="image" alt="SEDML" /></td>
<td>NuML?</td>
</tr>
<tr>
<td><strong>Terminologies</strong></td>
<td><img src="image" alt="SBO" /></td>
<td><img src="image" alt="KISAO" /></td>
<td><img src="image" alt="TEDDY" /></td>
</tr>
</tbody>
</table>

Is the matrix of standards complete?
Dimension 1: Covering the entire model life-cycle

<table>
<thead>
<tr>
<th>Model generation</th>
<th>Model structure</th>
<th>Parametrisation</th>
<th>Simulations and analysis</th>
<th>Numerical results</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>S8ML™</td>
<td>?</td>
<td>SED ML</td>
<td>NUML?</td>
</tr>
</tbody>
</table>
Dimension 2: Representing the levels of discourse

- Graphical representation
- Biological semantics
- Initial conditions (numbers)
- Model semantics (structure)