Report on the status of SBGN ER and proposed extensions
Graphs are everywhere in biology
Can-this be understood by biologists?
Can-this be understood by biologists?

Stimulates? but ... what exactly?

Associates into?

Translocates?

No idea. Reciprocal stimulation?

Stimulates gene transcription?

Is degraded?
Ambiguity of usual representations

$X \rightarrow Y$

is transformed into

translocates ($X \Rightarrow Y$)

is degraded into

associates into

dissociates into

stimulates the activity of

stimulates the expression of

catalyses the formation of

$X$ inhibits $Y$

inhibition
Kohn’s Molecular Interaction Maps

Kitano's Notation

**State node symbols**
- Protein
- Receptor
- Ion channel (closed)
- Ion channel (open)
- Truncated protein
- Gene
- RNA
- Anti-sense RNA
- Ion
- Simple molecule
- Unknown
- Phenotype
- Homodimer / N-mer with N stacked symbols
- Active protein

**Arc symbols**
- (Transit node and edges)
  - State transition
  - Known transition omitted
  - Unknown transition
  - Bidirectional transition
  - Translocation
  - Association
  - Dissociation
  - Truncation
  - Promote transition
  - Inhibit transition
  - Add reactant
  - Add product
  - AND
  - OR

**Reduced notation symbols**
- Category-I reduced notation
  - Degradation
  - Transcription
  - Translation
  - Module
- Category-II reduced notation (viewer only)
  - Activation/inhibition/modification
  - Index
- Node structure
  - Residue modification
  - Complex state node
  - Promoter and coding structure for gene
  - Exon structure for RNA

**References**
Enters
The Systems Biology Graphical Notation

http://www.sbgn.org/
What is SBGN?

- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
  Re-use existing symbols
  Smooth learning curve
- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models and growing software support
- Initiated by Hiroaki Kitano. Developed over four years by a diverse community
Graph trinity: three languages in one notation

Process Descriptions
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

Entity Relationships
- Unambiguous
- Mechanistic
- Non-sequential
- Independence of relationships

Activity Flows
- Ambiguous
- Conceptual
- Sequential

Three orthogonal projections of biology
The Systems Biology Graphical Notation

Nicolas Le Novère1, Michael Hucka2, Huaiyu Mi3, Stuart Moodie4, Falk Schreiber5,6, Anatoly Sorokin7, Emek Demir8, Katja Wegner9, Mirit I Aladjem10, Sarala M Wimalaratne11, Frank T Bergman12, Ralph Gauges13, Peter Ghazal4,14, Hideya Kawaji15, Lu Li1, Yukiko Matsuoka16, Alice Villéger17,18, Sarah E Boyd19, Laurence Calzone20, Melanie Courtot21, Ugur Dogrusoz22, Tom C Freeman14,23, Akira Funahashi24, Samik Ghosh16, Akiya Jouraku24, Sohyeong Kim10, Fedor Kolpakov25,26, Augustin Luna10, Sven Sahle13, Esther Schmidt1, Steven Watterson4,22, Guanming Wu27, Igor Goryanin4, Douglas B Kell18,28, Chris Sander8, Herbert Sauro12, Jacky L Snoep29, Kurt Kohn10 & Hiroaki Kitano16,30,31

1EMBL European Bioinformatics Institute, Hinxton, UK. 2Engineering and Applied Science, California Institute of Technology, Pasadena, California, USA. 3SRI International, Menlo Park, California, USA. 4Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh, UK. 5Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany. 6Institute of Computer Science, University of Halle, Halle, Germany. 7School of Informatics, University of Edinburgh, Edinburgh, UK. 8Memorial Sloan Kettering Cancer Center - Computational Biology Center, New York, NY, USA. 9Science and Technology Research Institute, University of Hertfordshire, Hatfield, UK. 10National Cancer Institute, Bethesda, Maryland, USA. 11Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand. 12Department of Bioengineering, University of Washington, Seattle, Washington, USA. 13BIOQUANT, University of Heidelberg, Heidelberg, Germany. 14Division of Pathway Medicine, University of Edinburgh Medical School, Edinburgh, UK. 15Riken OMICS Science Center, Yokohama City, Kanagawa, Japan. 16The Systems Biology Institute, Tokyo, Japan. 17School of Computer Science, University of Manchester, Manchester, UK. 18Manchester Interdisciplinary Biocentre, Manchester, UK. 19Clayton School of Information Technology, Faculty of Information Technology, Monash University, Melbourne, Victoria, Australia. 20U900 INSERM, Paris Mines Tech, Institut Curie, Paris, France. 21Terry Fox Laboratory, British Columbia Cancer Research Center, Vancouver, British Columbia, Canada. 22Bilkent Center for Bioinformatics, Bilkent University, Ankara, Turkey. 23The Roslin Institute, University of Edinburgh, Midlothian, UK. 24Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama, Japan. 25Institute of Systems Biology, Novosibirsk, Russia. 26Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia. 27Ontario Institute for Cancer Research, Toronto, Ontario, Canada. 28School of Chemistry, University of Manchester, Manchester, UK. 29Department of Biochemistry, Stellenbosch University, Matieland, South Africa. 30Sony Computer Science Laboratories, Tokyo, Japan. 31Okinawa Institute of Science and Technology, Okinawa, Japan. Correspondence should be addressed to N.L.N. (lenov@ebi.ac.uk).
Systems Biology Graphical Notation: Activity Flow language Level 1

Version 1.1
Date: October 6, 2010

Editors:
Nicolas Le Novère
EMBL European Bioinformatics Institute, UK
Stuart Moodie
CSIR, University of Edinburgh, UK
Falk Schreiber
JPK Göttingen & University of Bielefeld, Germany
Hans-Peter Meinhardt
University of Heidelberg, Germany

To discuss any aspect of SBGN, please send your messages to the mailing list sbgn-discuss@lists.sourceforge.net. To get subscribed to the mailing list or to contact us directly, please write to sbgn-admin@lists.sourceforge.net. Any reports and specific comments about the specification should be entered in the issue tracker: https://sourceforge.net/tracker/?atid=1079010.
Entity Relationships can be viewed as rules

If A exists, the assignment of the value P to the state variable T of B is increased
If A exists, the assignment of the value P to the state variable T of B is increased

(A stimulates the phosphorylation of B on the threonine)
Entity Relationships can be viewed as rules

If A exists, the assignment of the value P to the state variable T of B is increased.

If P is assigned to the state variable T of B, the assignment of the value P to the state variable S of C is decreased.
Multi-state and combinatorial explosion

Process Descriptions:
“once a state variable value, always a state variable value”

$2^{12} = 4096$ states
(i.e. EPN glyphs) for EGFR
and 4096 complexes between EGFR and targets
continuants, things that exist (or not)
**Entity Relationships L1 V1 reference card**

**Entity Nodes**

- **Interactors**
  - LABEL: entity
  - ⬤: outcome

- **LABEL**: perturbing agent

**Logical Operators**

- **AND**: and operator
- **OR**: or operator
- **NOT**: not operator
- τ: delay operator

**Auxiliary units**

- pre:label: unit of information
- variable: state variable
- ⬤: existence
- ⬤: location

**Reference nodes**

- LABEL: annotation

**Relationship Nodes**

**Statements**

- N: interaction
- →: assignment
- ←: observable

**Influence**

- ▲: modulation
- ▲: stimulation
- ▲: necessary stimulation
- ▼: absolute stimulation
- ▼: inhibition
- ▼: absolute inhibition
- ▼: logic arc

*occurrents, events that may happen (or not)*
## Entity Relationships L1 V1 syntax

<table>
<thead>
<tr>
<th>symbols \ Arc</th>
<th>interaction</th>
<th>modulation</th>
<th>stimulation</th>
<th>inhibition</th>
<th>necessary stimulation</th>
<th>absolute stimulation</th>
<th>absolute inhibition</th>
<th>logic arc</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>IO</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>outcome</td>
<td>I(1)O(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
</tr>
<tr>
<td>and</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)O</td>
</tr>
<tr>
<td>or</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)O</td>
</tr>
<tr>
<td>not</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)O(1)</td>
</tr>
<tr>
<td>delay</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)</td>
<td>I(1)O(1)</td>
</tr>
<tr>
<td>perturbing agent</td>
<td></td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>unit of information</td>
<td></td>
<td>IO</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>state variable</td>
<td>I(1)O(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>modulation</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>stimulation</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>inhibition</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>necessary stimulation</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>absolute stimulation</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>absolute inhibition</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>assignment</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>interaction</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>phenotype</td>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>
3.4.2 Influences

A modulation (Section 2.4.3.1) linking an entity node $E$ and a relationship $R$ means: “If $E$ exists then $R$ is either reinforced or weakened”.

A stimulation (Section 2.4.3.2) linking an entity node $E$ and a relationship $R$ means: “If $E$ exists then $R$ is reinforced” or “If $E$ exists then the probability of $R$ is increased”.

An absolute stimulation (Section 2.4.3.6) linking an entity node $E$ and a relationship $R$ means: “If $E$ exists then $R$ always takes place”.

A necessary stimulation (Section 2.4.3.4) linking an entity node $E$ and a relationship $R$ means: “$R$ only takes place if $E$ exists.

An inhibition (Section 2.4.3.3) linking an entity node $E$ and a relationship $R$ means: “If $E$ exists then $R$ is weakened” or “If $E$ exists then the probability of $R$ is lowered”.

An absolute inhibition (Section 2.4.3.5) linking an entity node $E$ and a relationship $R$ means: “If $E$ exists then $R$ never takes place”. 
ER map of calcium-regulated synaptic plasticity

increases synaptic weight

decreases synaptic weight
Entity Relationships L1 V2

nested entities
A and B are part of X
X interacts with Y

domainA

domainB

entityX

tentityY
A of X interacts with Y
Translocation of X in the nucleus

The value “nucleus” is assigned to the “location” of entityX, domainA and domainB
The value “false” is assigned to the “existence” of entityX, domainA and domainB
The value “membrane” is assigned to the “location” of domainB but not to domainB and entityX
The value “false” is assigned to the “existence” of domainA but not to domainA and entityX that remain true.
Same molecule => confusing
T306 part of both low and high affinity => confusing
Pending issues

• Logical operation on statements: In L1 V1, logical operators only output influences. Implicit XOR on variable assignments. Nothing for interactions.

• Topology of domains: Might be necessary if we want adoption by genomics, synthetic biology etc.

• Identification of instances: How to differentiate between several instances of the same entity, differentially involved in a relationships (e.g. trans-phosphorylation)?

• Identification of generics: How to lump together several entities for a given relationships? (e.g. MAPK instead of ERK1 and ERK2).
either C or the complex of A and B inhibits ...

The absence of T stimulates ...

U and V are necessary to the existence of ...
Implicit OR

Only ONE arc, not a superposition of two.
Logical operator participating to statements

- Current situation for interactions
- Proposal for interactions
- Proposal for assignment
Consequence: outcomes on influences

- value2 is sent to variable
- value1 is sent to variable
- variable has value1 or value2
- logic arc
- assignment

On all influences ...

A stimulates the assignment of value to variable

value is assigned to variable
Topology of domains
Generics in ER

ERK1 → P → MEK1 → P → MEK2 → P

ERK2 → P → "="

MSK1

ERK