

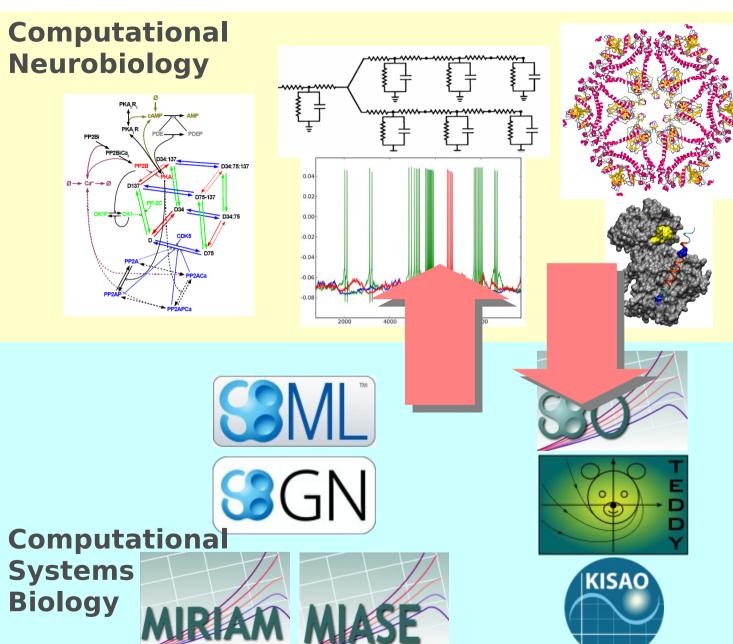
Pink Seminar

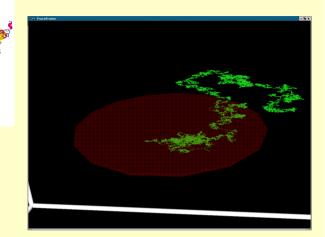
Nicolas Le Novère, EMBL-EBI





compneur group themes and projects









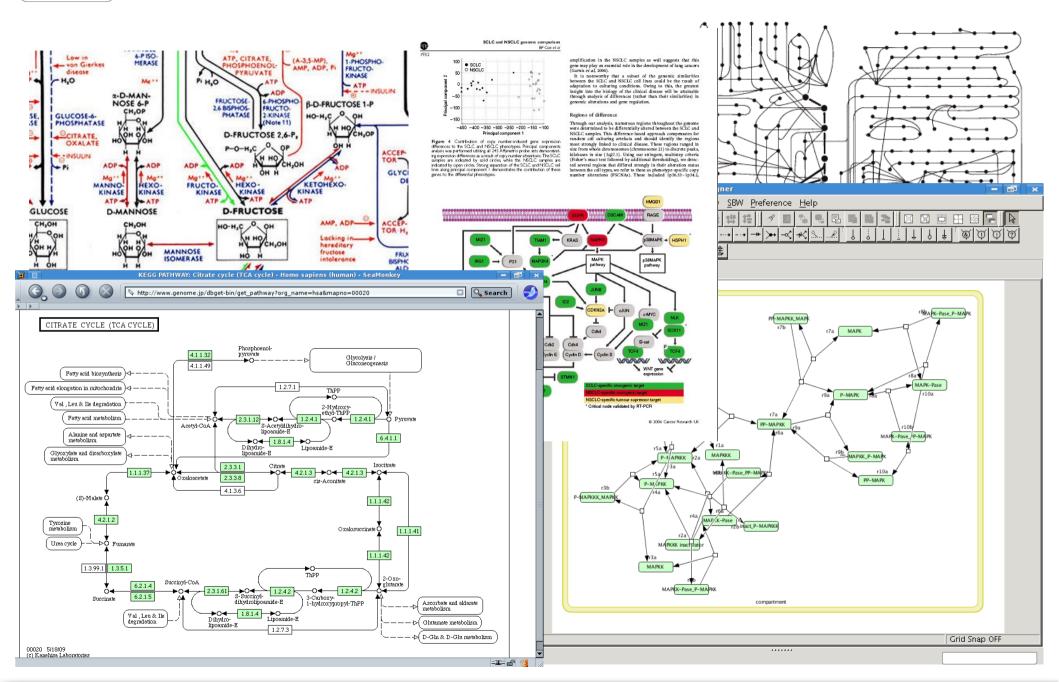


A common layout for the Tower of Babel: The Systems Biology Graphical Notation

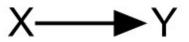
(on the behalf of SBGN editors, authors and contributors)



Graphs are everywhere in biology











is transformed into

translocates (X "=" Y)

is degraded into

associates into

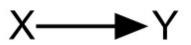
dissociates into

stimulates the activity of

stimulates the expression of

catalyses the formation of





X inhibits Y

is transformed into

translocates (X "=" Y)

is degraded into

associates into

dissociates into

stimulates the activity of

stimulates the expression of

catalyses the formation of





is transformed into

translocates (X "=" Y)

is degraded into

associates into

dissociates into

stimulates the activity of

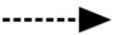
stimulates the expression of

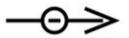
catalyses the formation of

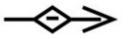














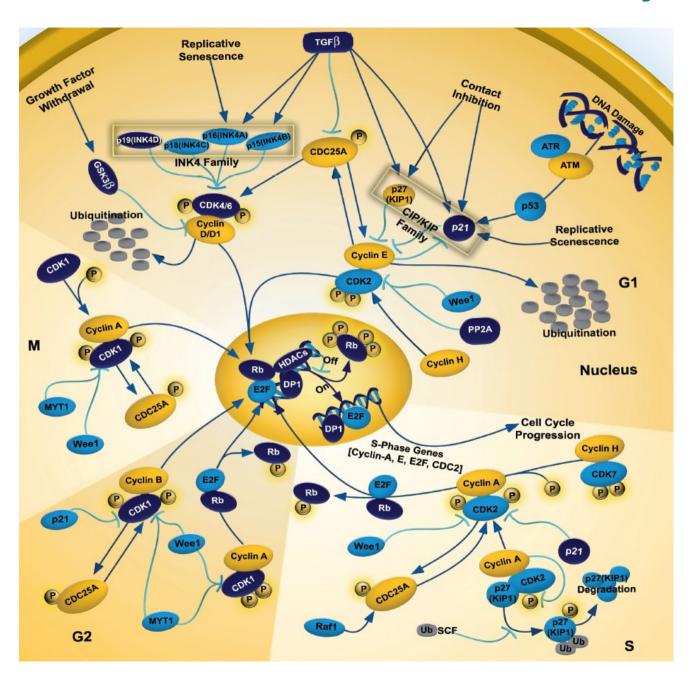






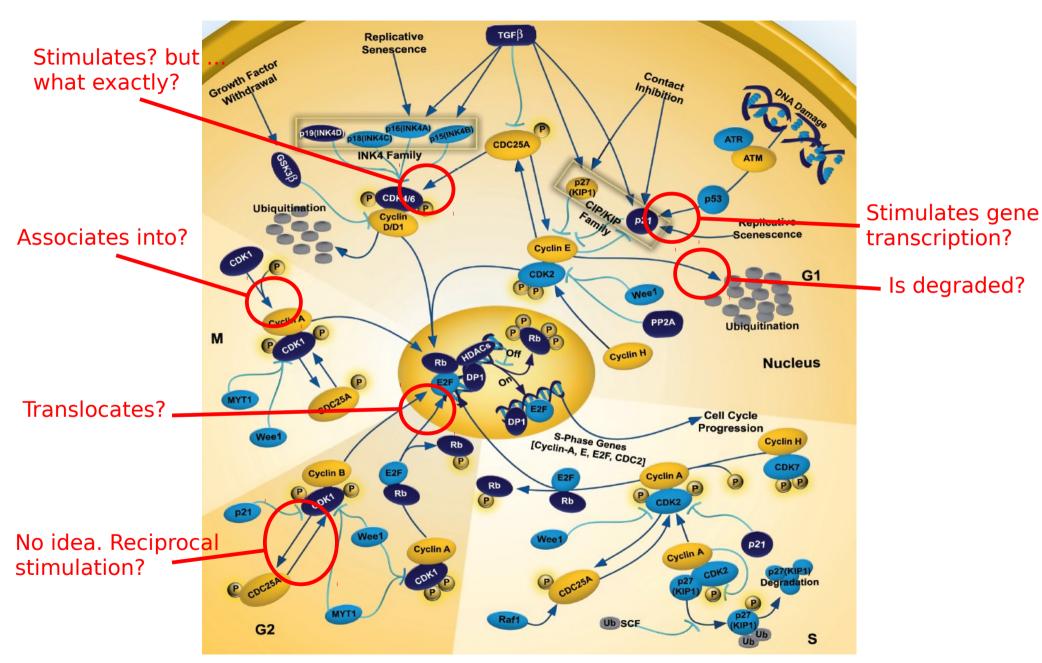


Can-this be understood by biologists?



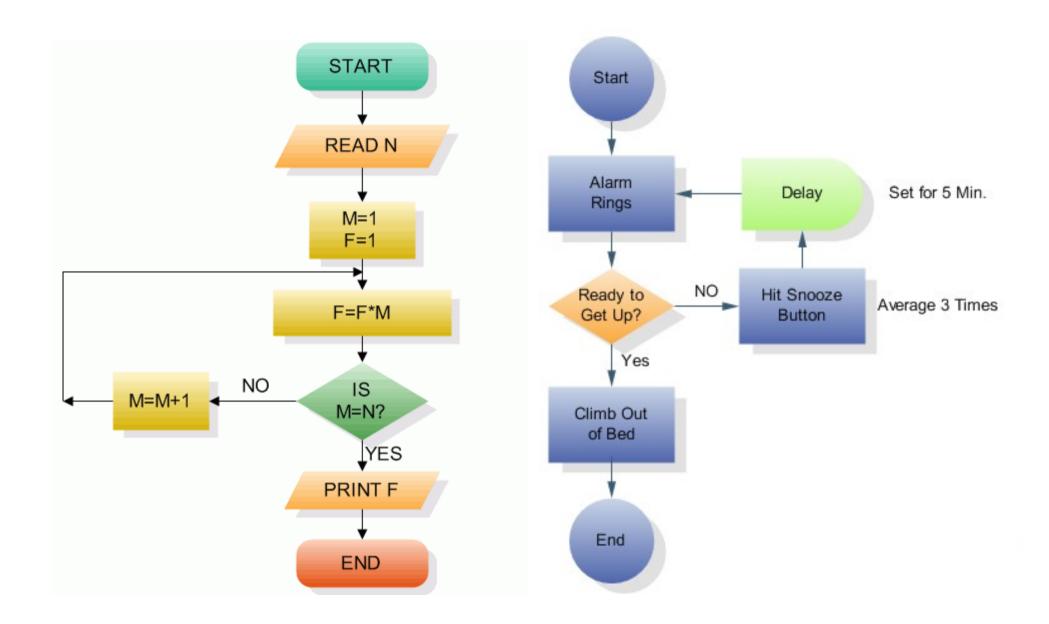


Can-this be understood by biologists?



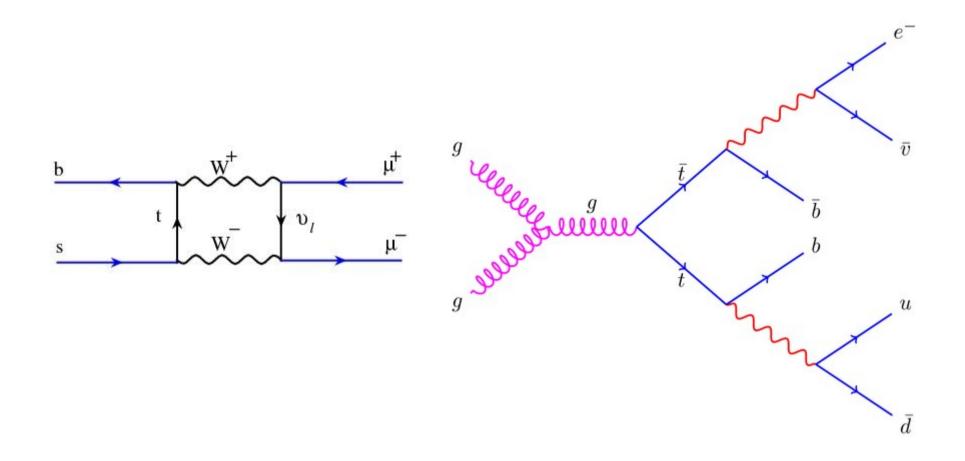


Every computer scientist understands those



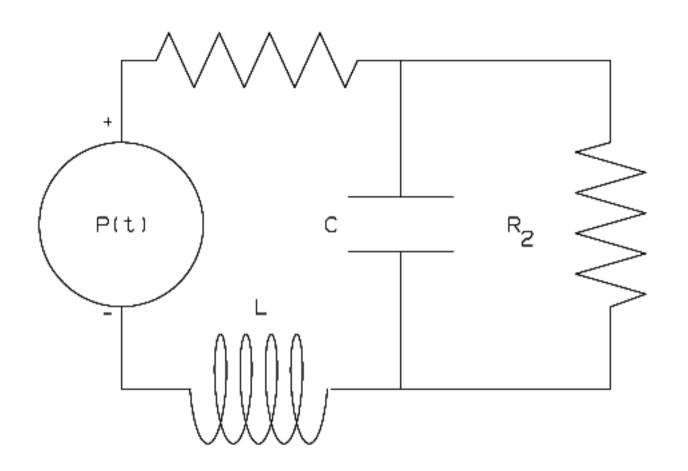


Every physicist understands those



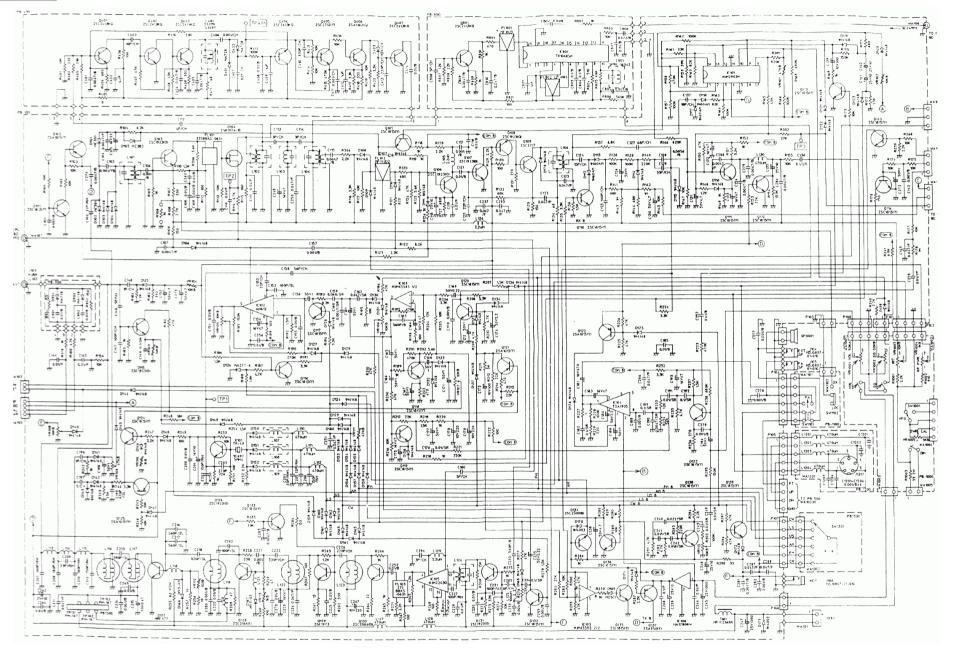


Every engineer understands that





Or that as a matter of fact





What did-those diagram bring?













What do-we expect in modern (future) life science



Basic science

Systems of Life
Systems Biology



Technology





An Introduction to Synthetic Biology

lanuary 2007





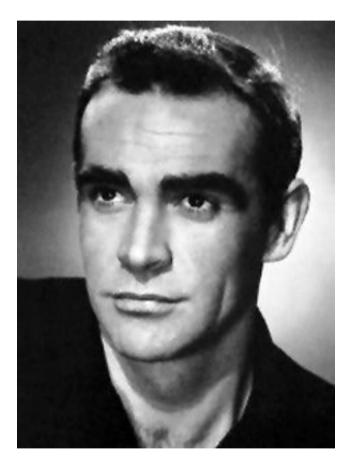
What happens if one cannot read the blueprint

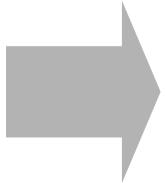






What happens if one cannot read the blueprint







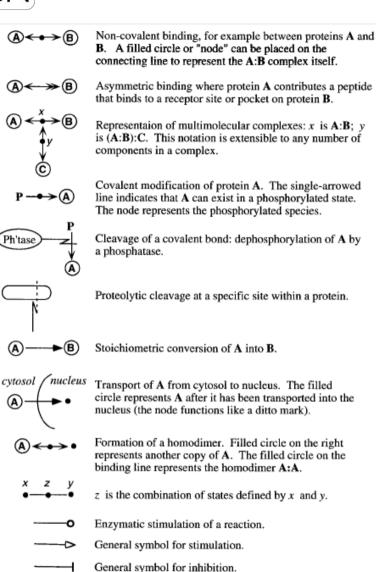


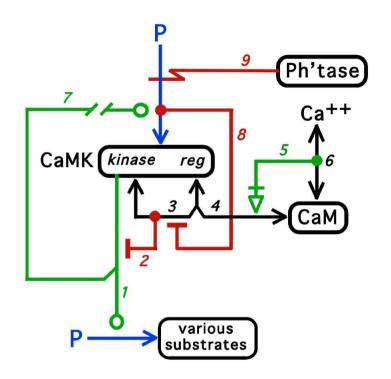
Why can-we understand electric diagrams?

- Standard symbols
 - Simple shapes, easily recognisable
 - Limited number of basic symbols (<70)</p>
 - Similarity of shapes reflects similarity of functions
- Unambiguous interpretation of the circuits
- Endorsed by the community for practical reasons
 - End-users: manufacturers
 - Tool developers
 - Publishing industry
 - Teaching communities



Kohn's Molecular Interaction Maps





- Kurt W Kohn (1998)
 Oncogene, 16: 1065-1075
- Kurt W. Kohn (1999)
 Mol Biol Cell,
 10(8):2703-2734

Degradation products

Shorthand symbol for transcriptional activation. Shorthand symbol for transcriptional inhibition.



Cell Cycle in MIM

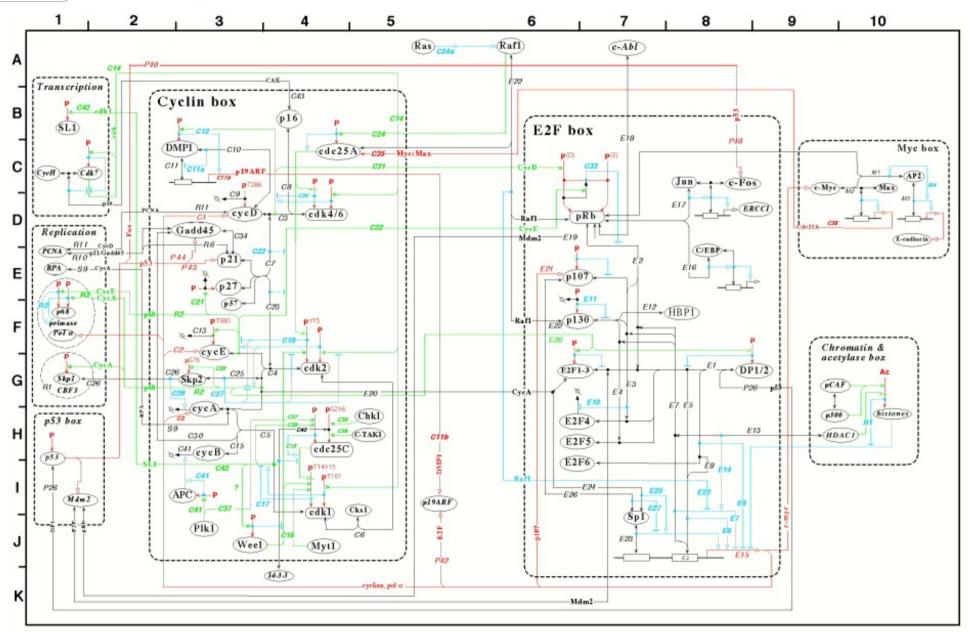
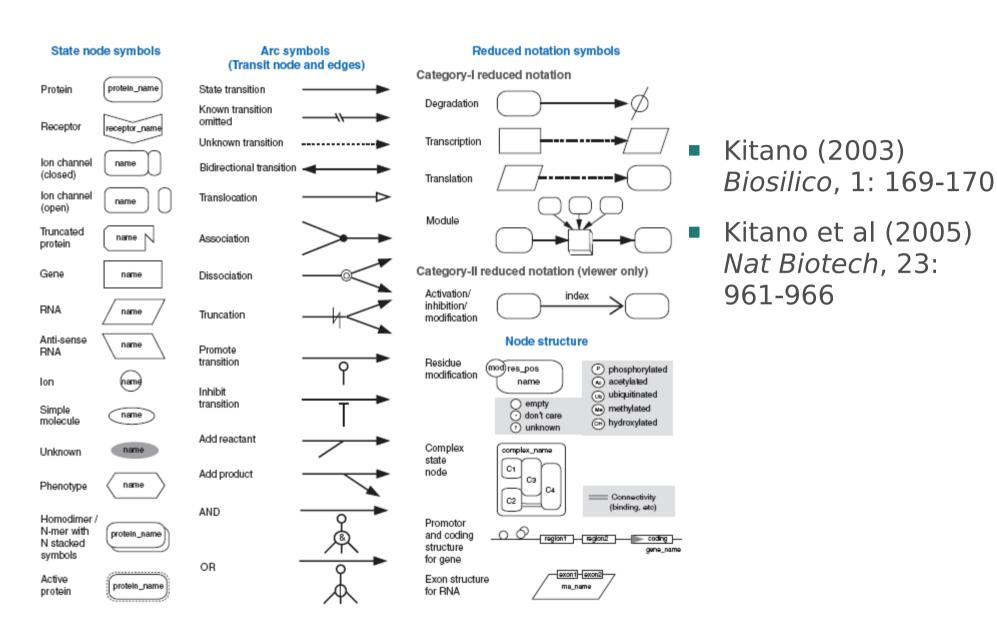


Figure 6A: The Cyclin - E2F cell cycle control system (version 3a - June 8, 1999)

MIM CAR I

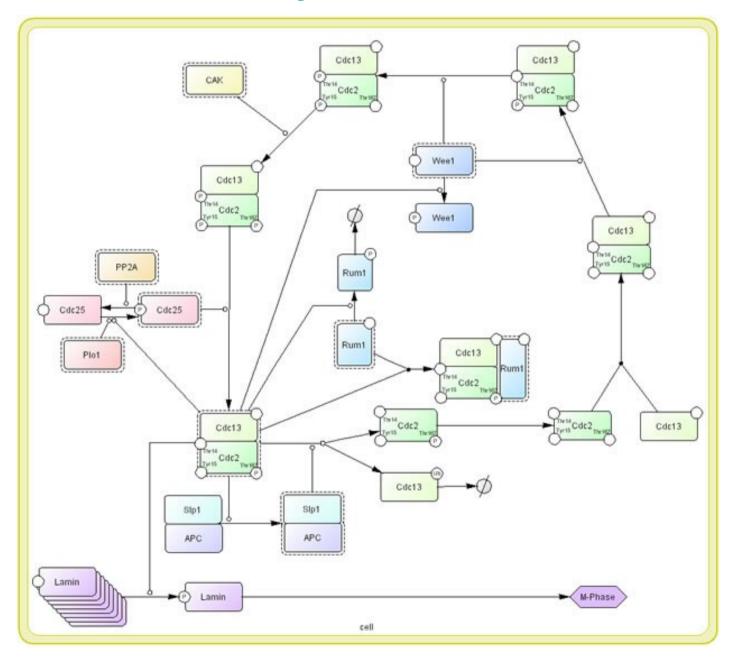


Kitano's Notation





Cell Cycle in Kitano's Process Diagram





Why those pioneer efforts did not make it

- Somehow fuzzy semantics
 - No structured data model or ontology behind the notation
 - Overlapping concepts rather than sub-classing
 - Gaps in the coverage of biochemistry or modelling
 - Ambiguous interpretation of the graph
- Little software support (except CellDesigner for Kitano's Process Diagrams)
- No community involvement
 - No systematic bug tracking and consistency checking
 - No comprehensive coverage (focussed on some use-cases)
 - No endorsement by the tool developers or by the end-users



Enters The Systems Biology Graphical Notation



http://www.sbgn.org/





- 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



The Systems Biology Graphical Notation

Nicolas Le Novère¹, Michael Hucka², Huaiyu Mi³, Stuart Moodie⁴, Falk Schreiber^{5,6}, Anatoly Sorokin⁷, Emek Demir⁸, Katja Wegner⁹, Mirit I Aladjem¹⁰, Sarala M Wimalaratne¹¹, Frank T Bergman¹², Ralph Gauges¹³, Peter Ghazal^{4,14}, Hideya Kawaji¹⁵, Lu Li¹, Yukiko Matsuoka¹⁶, Alice Villéger^{17,18}, Sarah E Boyd¹⁹, Laurence Calzone²⁰, Melanie Courtot²¹, Ugur Dogrusoz²², Tom C Freeman^{14,23}, Akira Funahashi²⁴, Samik Ghosh¹⁶, Akiya Jouraku²⁴, Sohyoung Kim¹⁰, Fedor Kolpakov^{25,26}, Augustin Luna¹⁰, Sven Sahle¹³, Esther Schmidt¹, Steven Watterson^{4,22}, Guanming Wu²⁷, Igor Goryanin⁴, Douglas B Kell^{18,28}, Chris Sander⁸, Herbert Sauro¹², Jacky L Snoep²⁹, Kurt Kohn¹⁰ & Hiroaki Kitano^{16,30,31}

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. 5 Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany. 6 Institute 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. 14 Division of Pathway Medicine, University of Edinburgh Medical School, Edinburgh, UK. 15 Riken OMICS Science Center, Yokohama City, Kanagawa, Japan. 16 The Systems Biology Institute, Tokyo, Japan. 17 School 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. 20 U900 INSERM, Paris Mines Tech, Institut Curie, Paris, France. 21 Terry Fox Laboratory, British Columbia Cancer Research Center, Vancouver, British Columbia, Canada. 22 Bilkent Center for Bioinformatics, Bilkent University, Ankara, Turkey. 23 The Roslin Institute, University of Edinburgh, Midlothian, UK. 24 Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama, Japan. 25 Institute of Systems Biology, Novosibirsk, Russia. 26 Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia. 27 Ontario Institute for Cancer Research, Toronto, Ontario, Canada, 28 School of Chemistry, University of Manchester, Manchester, UK. 29 Department of Biochemistry, Stellenbosch University, Matieland, South Africa. 30 Sony Computer Science Laboratories, Tokyo, Japan. 31 Okinawa Institute of Science and Technology, Okinawa, Japan. Correspondence should be addressed to N.L.N. (lenov@ebi.ac.uk).

NATURE BIOTECHNOLOGY VOLUME 27 NUMBER 8 AUGUST 2009

741

39 authors, 31 affiliations



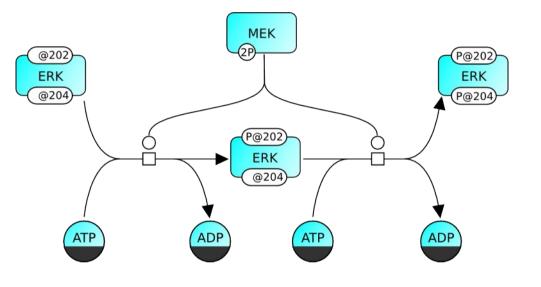


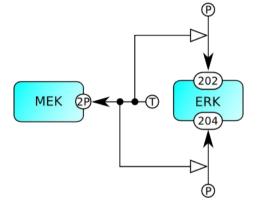
Graph trinity: three languages in one notation

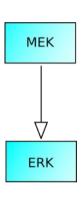
<u>Process Descriptions</u>

Entity Relationships

Activity Flows







- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

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

- Ambiguous
- Conceptual
- Sequential

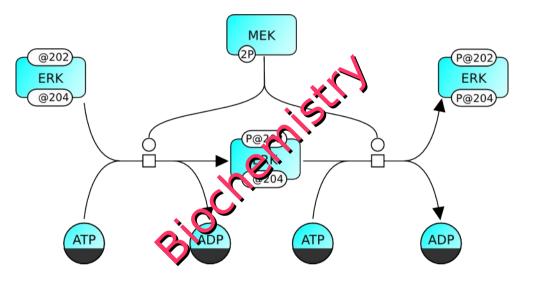


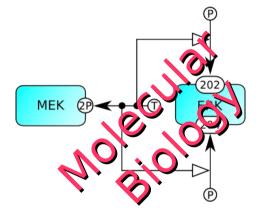
Graph trinity: three languages in one notation

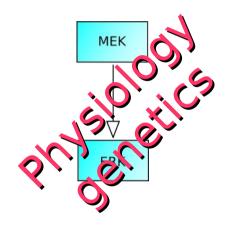
Process Descriptions

Entity Relationships

Activity Flows







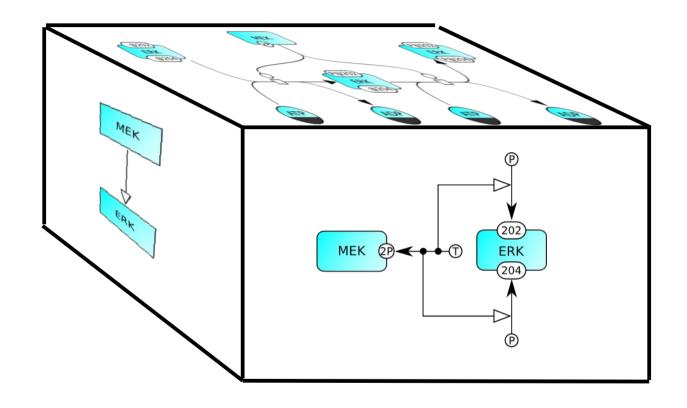
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

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

- Ambiguous
- Conceptual
- Sequential

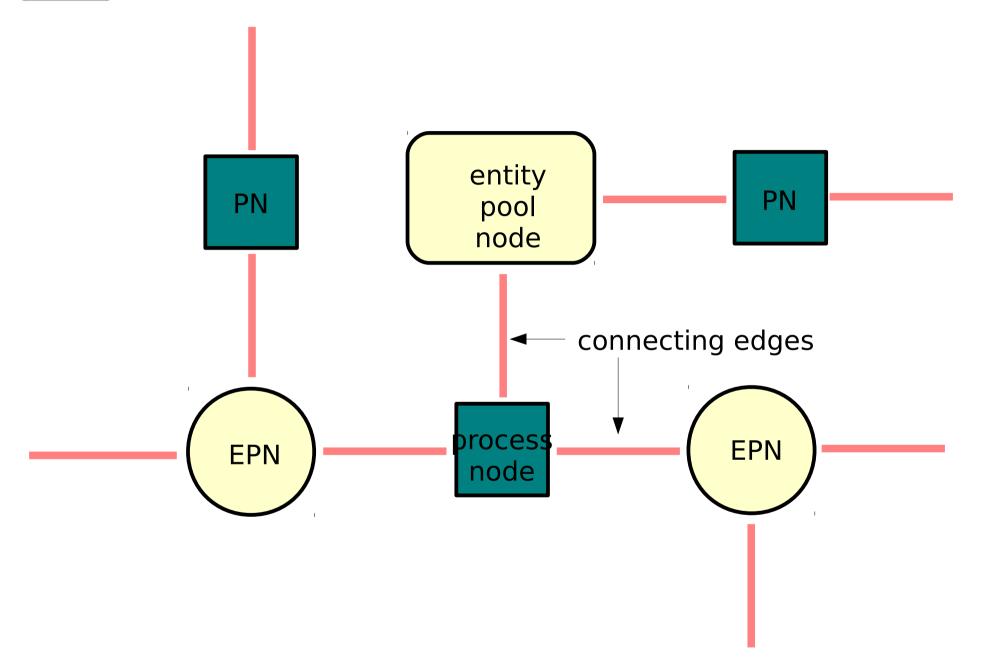


Three orthogonal projections of biology



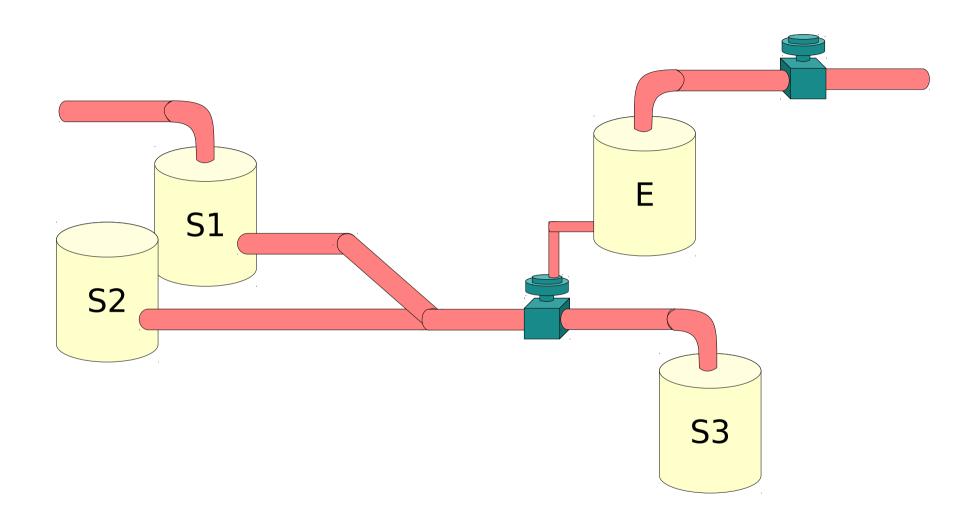


Process Descriptions are bipartite graphs



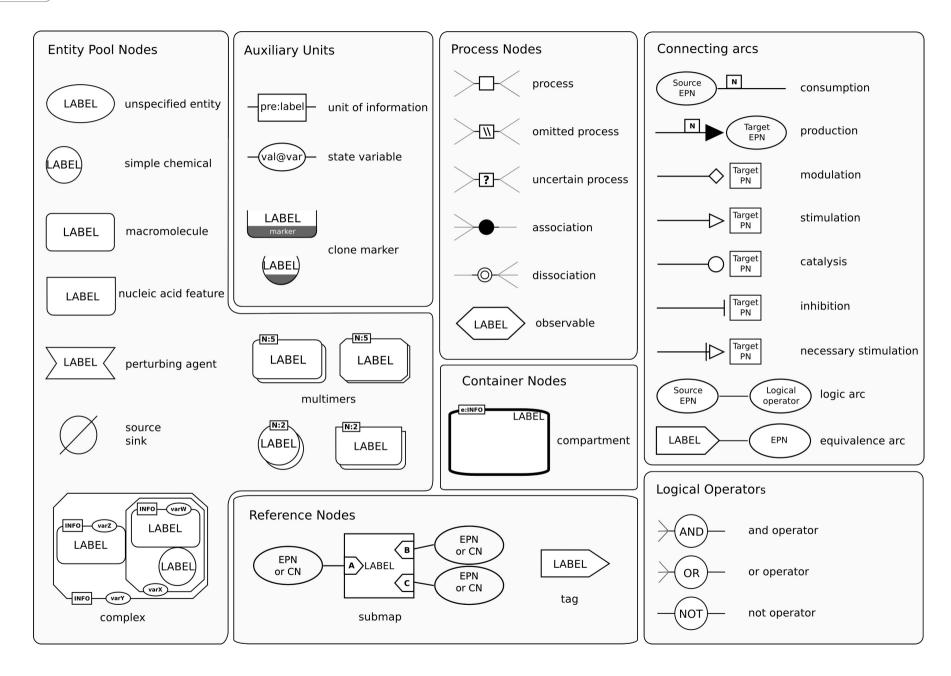


Process Descriptions can be viewed as pipelines



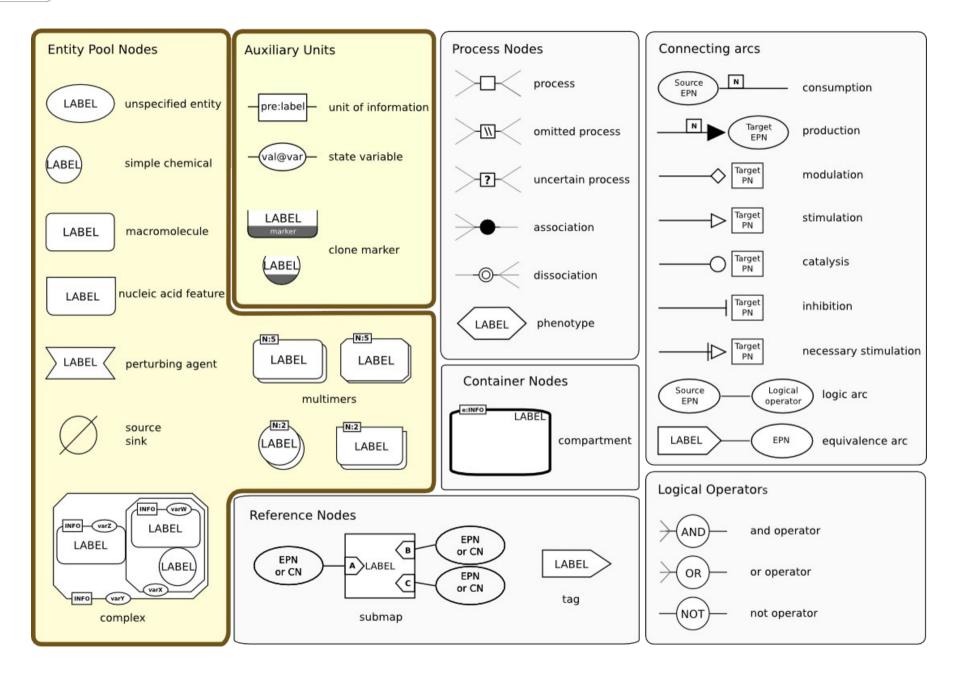


SBGN Process Description L1 V1.1 reference card



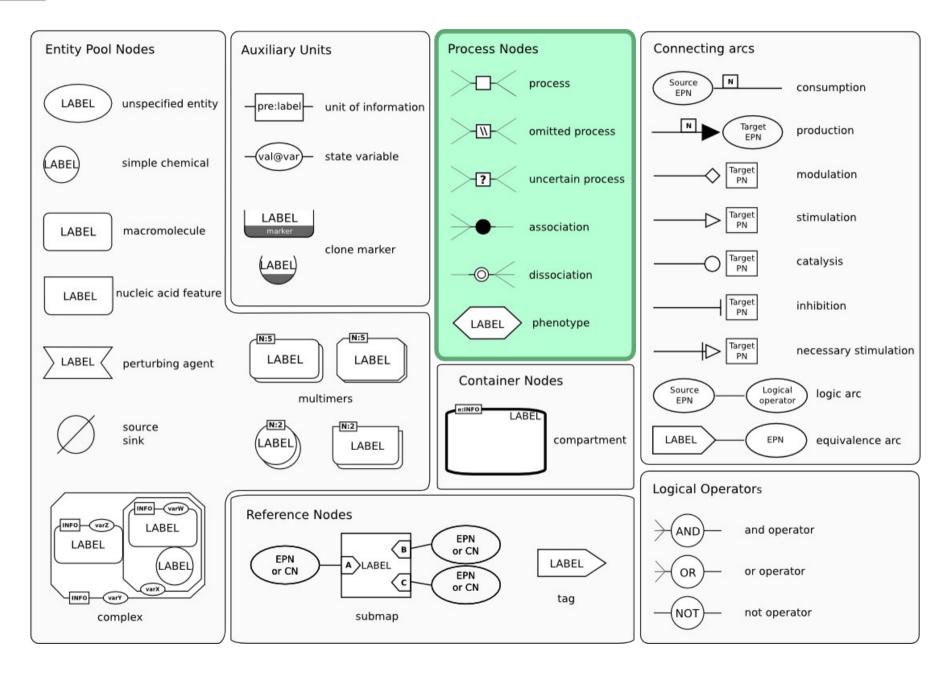


Entity Pool Nodes



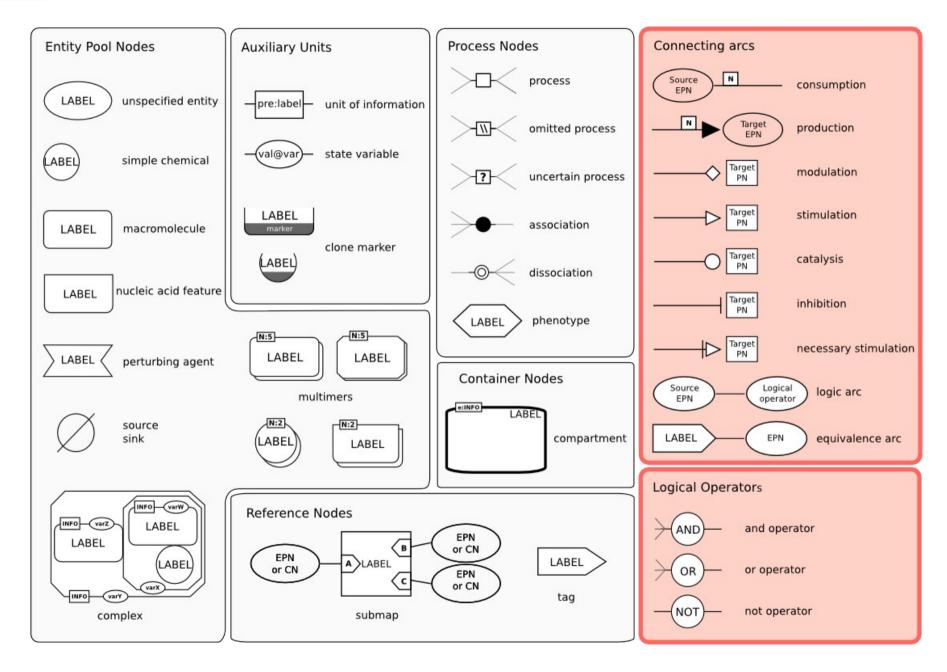


Process Nodes



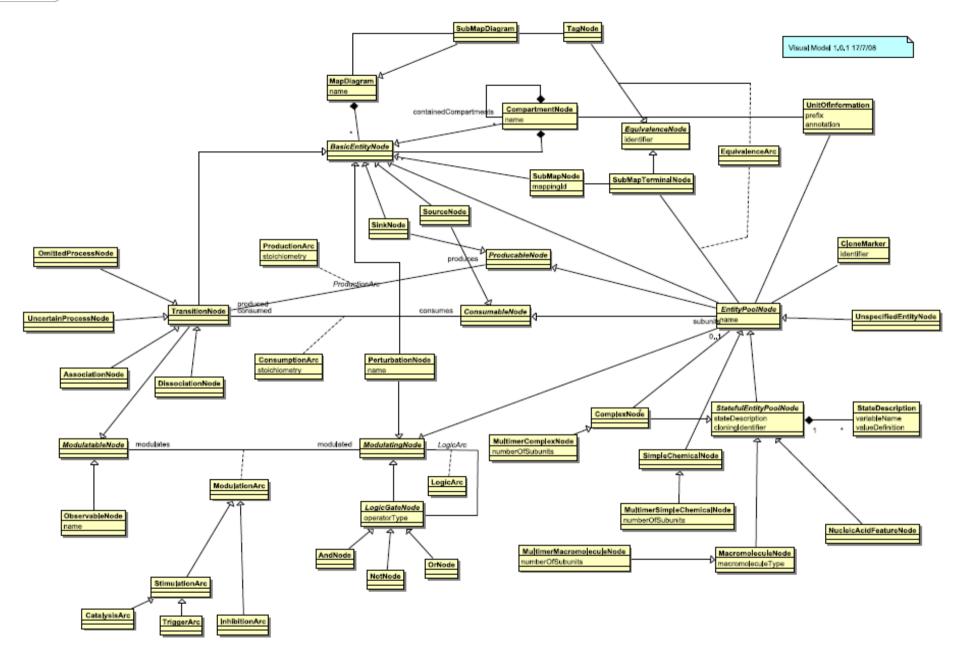


Edges





Process Descriptions data model





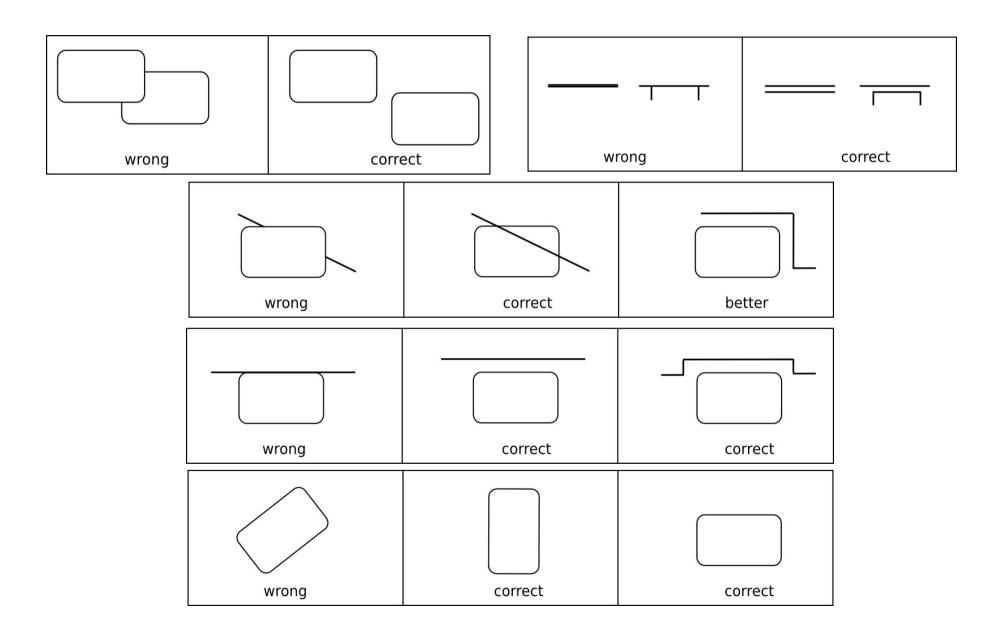
Process Descriptions syntax definition

$Arc \backslash EPN$	macromolecule	simple chemical	unspecified entity	multimer	complex	nucleic acid feature	tag	source/sink	perturbation	observable	submap
consumption	Ι	Ι	Ι	Ι	Ι	Ι		Ι			
production	О	О	О	О	О	О		О			
modulation	Ι	Ι	Ι	Ι	Ι	Ι			Ι	О	
stimulation	Ι	Ι	Ι	Ι	Ι	Ι			Ι	О	
catalysis	Ι	Ι	Ι	Ι	Ι				Ι	О	
inhibition	Ι	Ι	Ι	Ι	Ι	Ι			Ι	О	
trigger	Ι	Ι	Ι	Ι	Ι	Ι			Ι	О	
logic arc	Ι	Ι	Ι	Ι	Ι	Ι					
equivalence arc	Ι	Ι	Ι	Ι	Ι	Ι	О				О

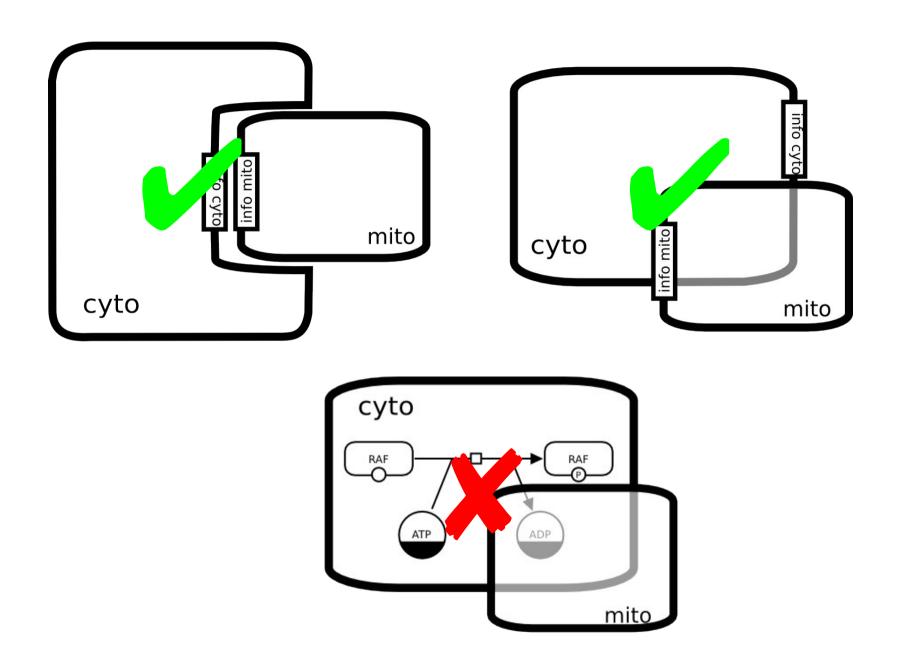
$Arc \backslash PN$	transition	omitted process	uncertain process	association	dissociation	and	or	not
consumption	О	О	О	О	O(1)			
production	Ι	Ι	Ι	I(1)	I			
modulation	О	О	О			I(1)	I(1)	I(1)
stimulation	О	О	О			I(1)	I(1)	I(1)
catalysis	О	О	О			I(1)	I(1)	I(1)
inhibition	О	О	О			I(1)	I(1)	I(1)
trigger	О	О	О			I(1)	I(1)	I(1)
logic arc						О	О	O(1)
equivalence arc								



Layout constraints

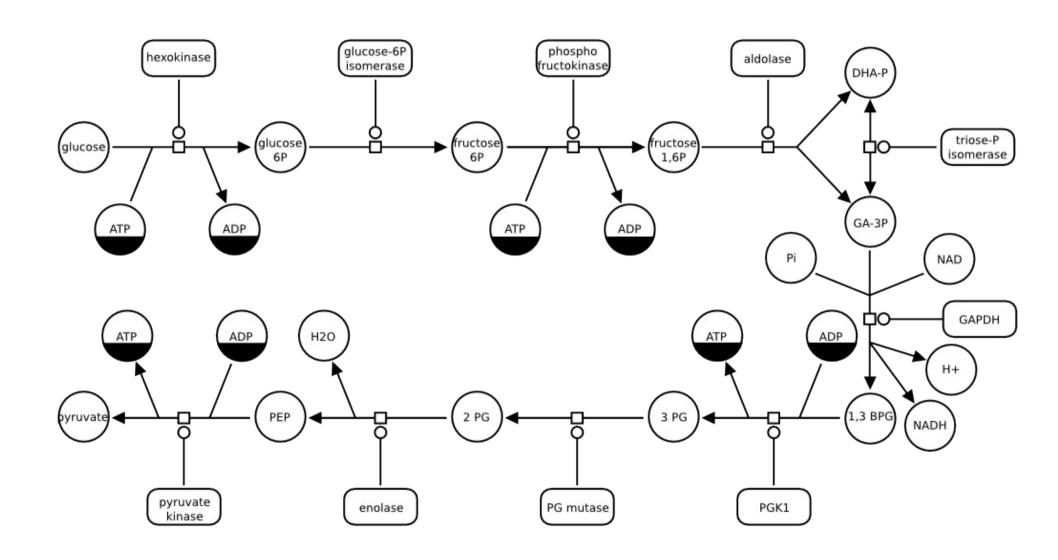






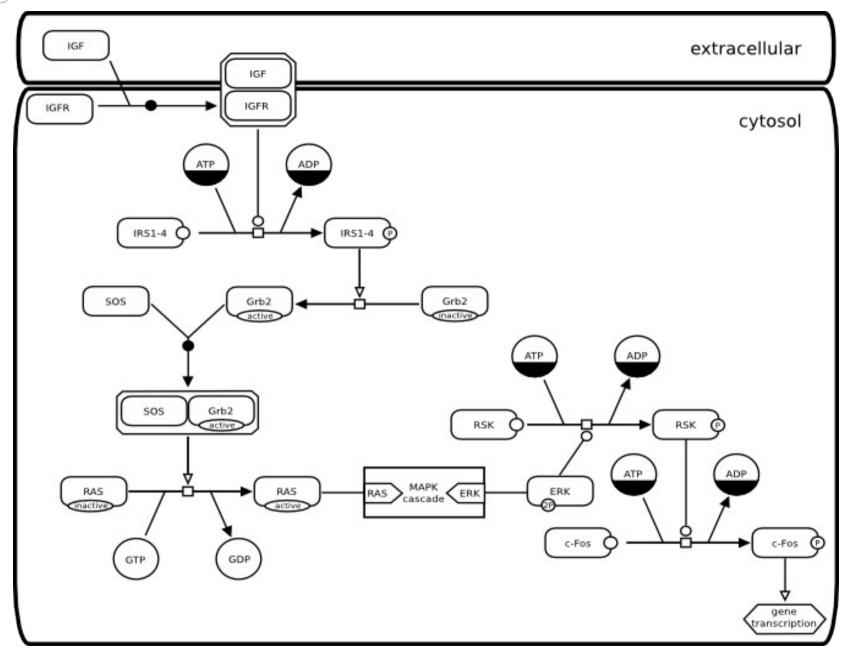


Metabolic network



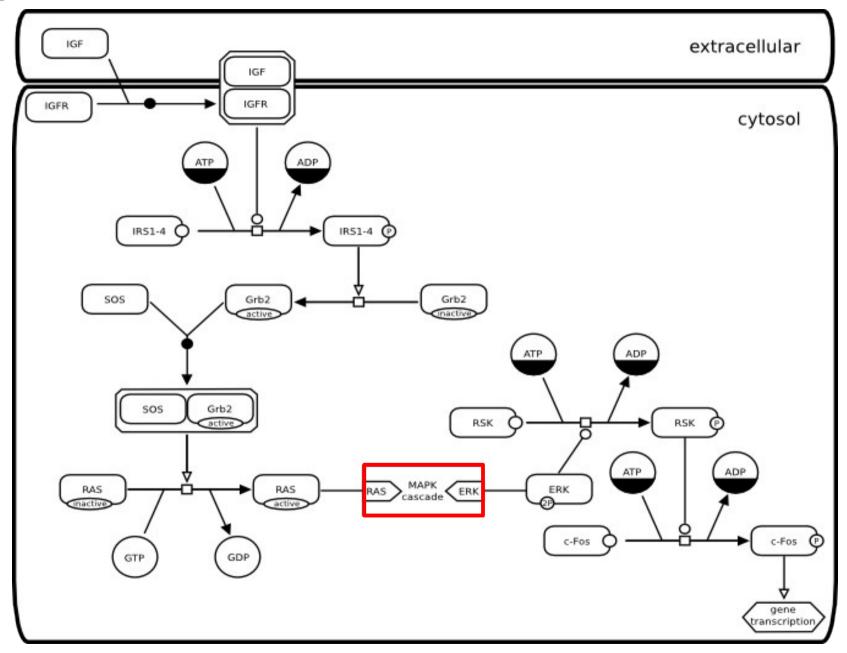


Signalling pathways



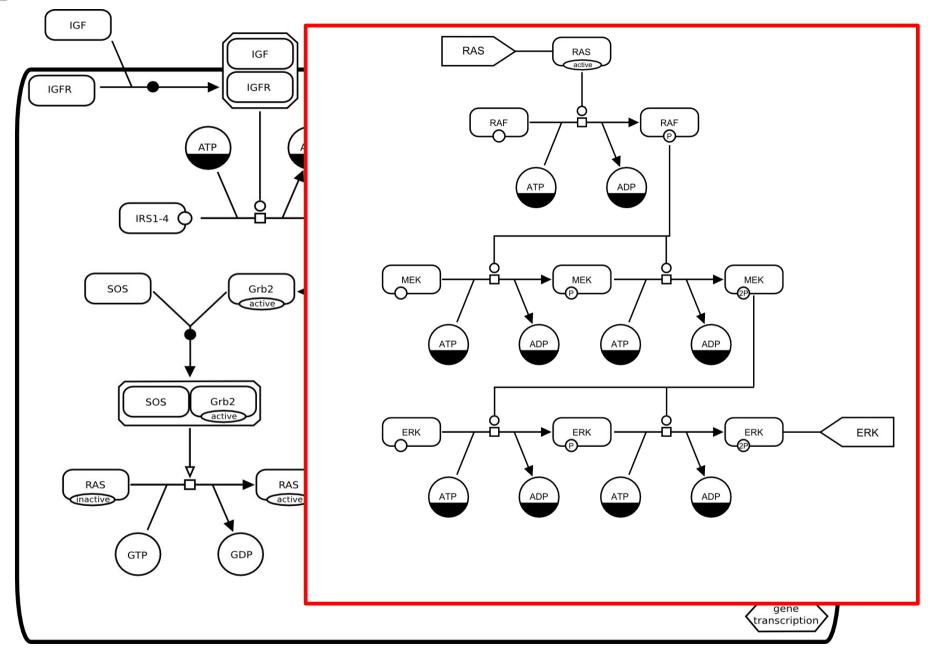


Submaps

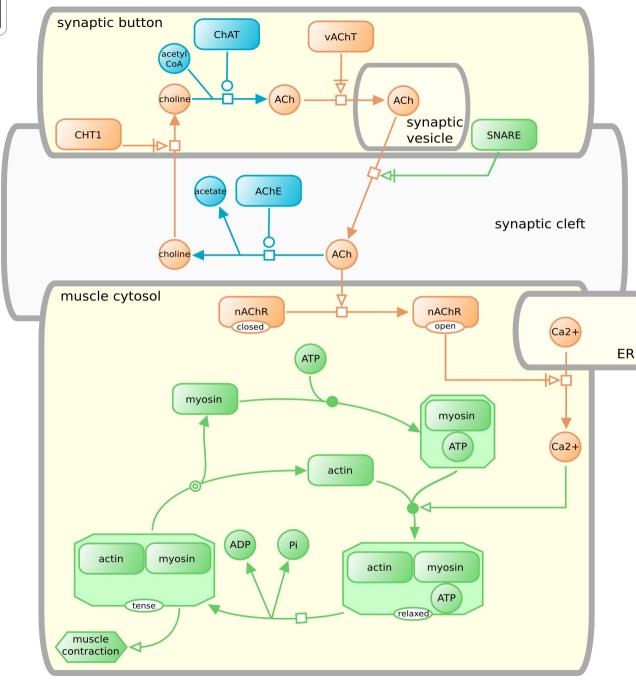




Submaps







multi-cellular processes

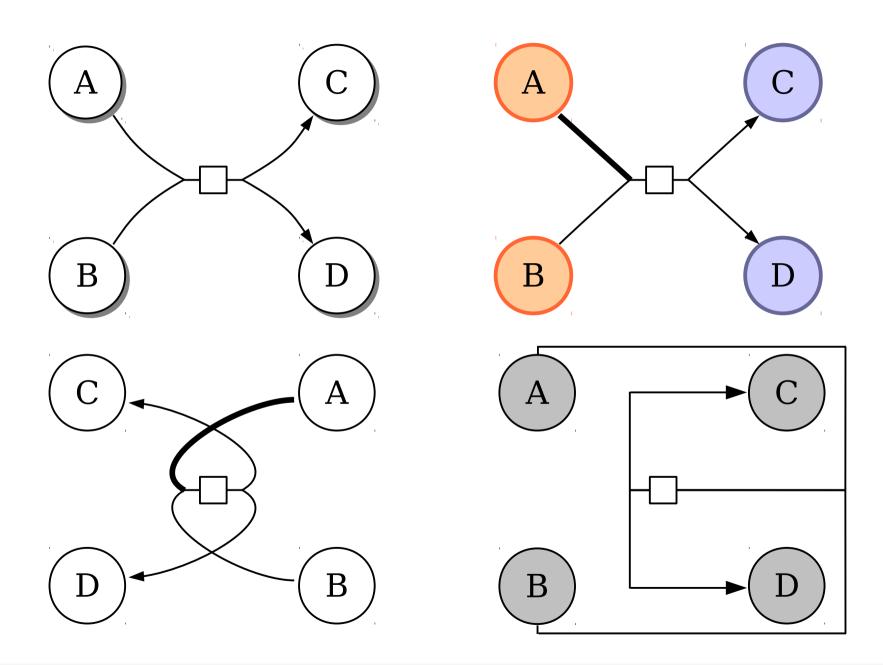
catalytic processes

transport processes

contractile proteins



All those diagrams are identical

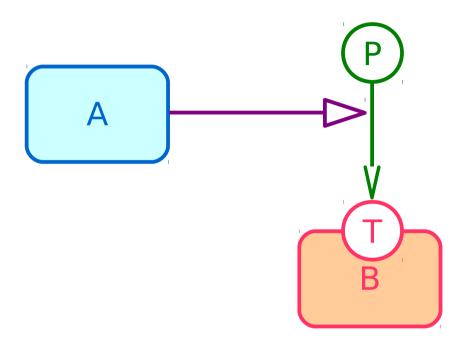




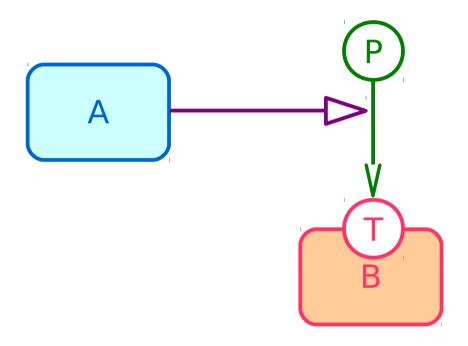
Pending issues

- Generics, i.e, entity pools representing several possible biochemical types. (e.g. MAPK instead of ERK1 and ERK2).
- Trans-compartment (e.g. transmembrane) structures
- Logical combination of state-variable values (and close-world/open-world position)
- Moving and transforming compartments
- Non-chemical entity pool nodes ("voltage", "pH" ...)



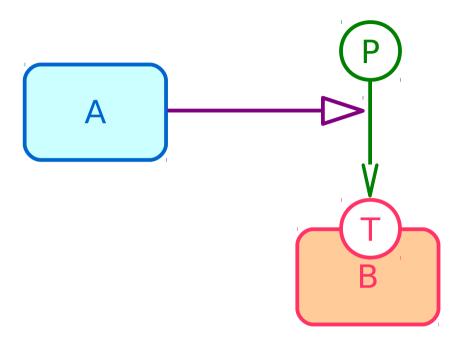






If A exists, the assignment of the value P to the state variable T of B is increased

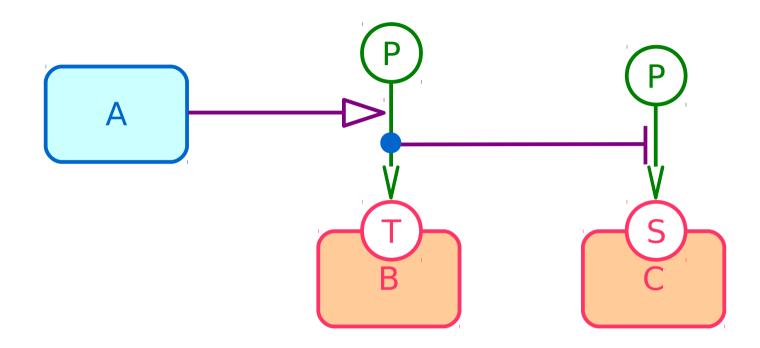




If A exists, the assignment of the value P to the state variable T of B is increased

(A stimulates the phosphorylation of B on the threonine)



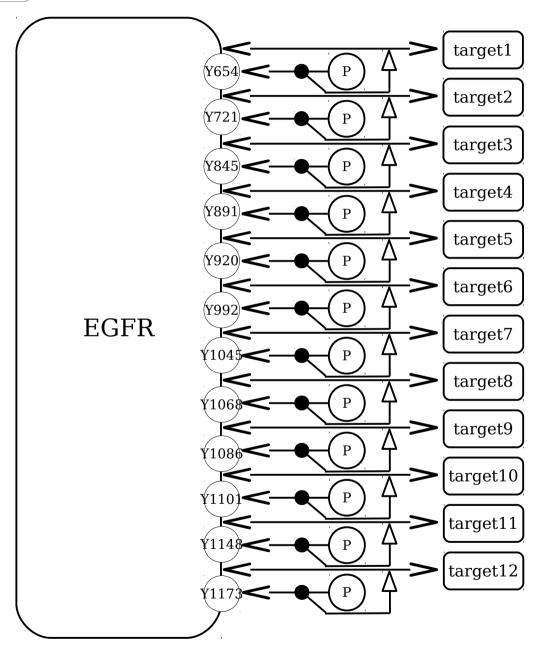


If A exists, the assignment of the value P to the state variable T of B is increased

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



Multistate and combinatorial explosion



Process Diagram:

"once a state variable value,
always a state variable value"

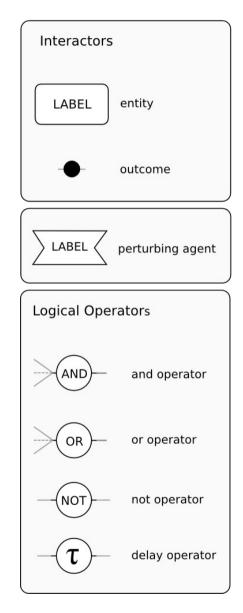
2¹² = 4096 states (i.e. EPN glyphs) for EGFR and 4096 complexes between EGFR and targets

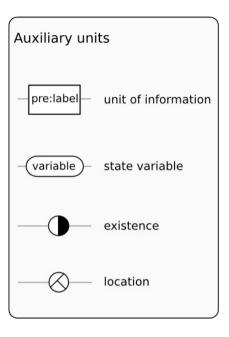


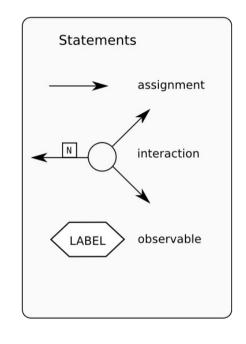
SBGN Entity Relationships L1 reference card

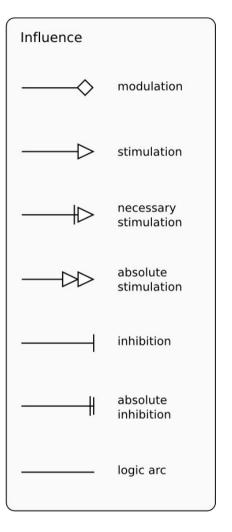
Entity Nodes

Relationship Nodes

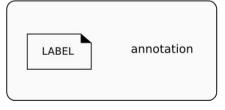










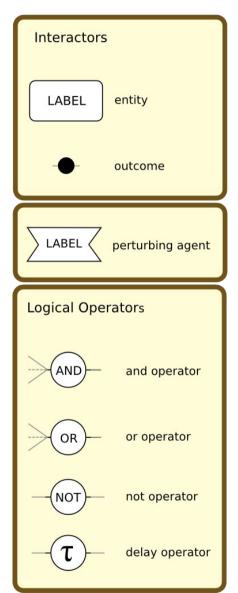


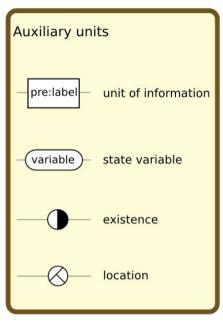


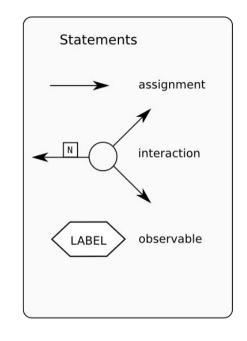
SBGN Entity Relationships L1 reference card

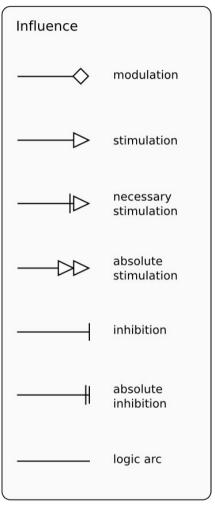
Entity Nodes

Relationship Nodes

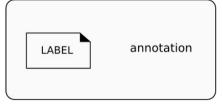








reference nodes



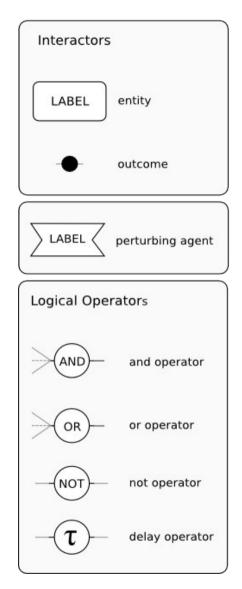
continuants, things that exists (or not)

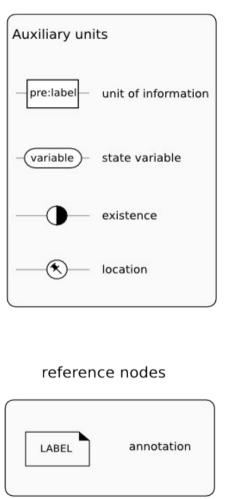


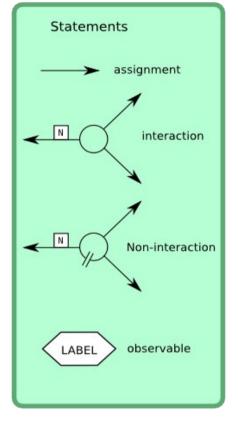
SBGN Entity Relationships L1 reference card

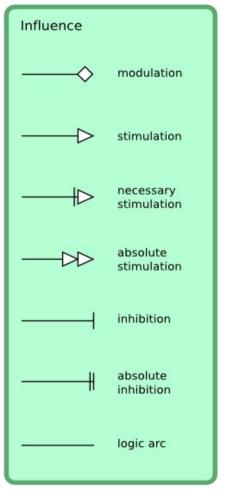
Entity Nodes

Relationship Nodes









occurrents, events that may happen (or not)



Entity Relationships syntax

symbols \ Arc	assignment	interaction	modulation	stimulation	inhibition	necessary stimulation	absolute stimulation	absolute inhibition	logic arc
entity		IO	I	Ι	I	I	Ι	I	I
outcome		I(1)O(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)
and			I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)O
or			I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)O
not			I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)O(1)
delay			I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)O(1)
perturbing agent			I	I	I	I	I	I	I
unit of information		IO							
state variable	I(1)O(1)								
modulation				0	0	О	0	О	
stimulation				0	0	О	0	0	
inhibition				0	0	О	0	0	
necessary stimulation				0	0	О	0	О	
absolute stimulation				0	О	О	О	О	
absolute inhibition				0	О	О	О	О	
assignment				0	0	О	0	О	
interaction				0	О	О	О	О	
phenotype				0	О	О	О	О	



Example of Entity Relationships semantics

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".

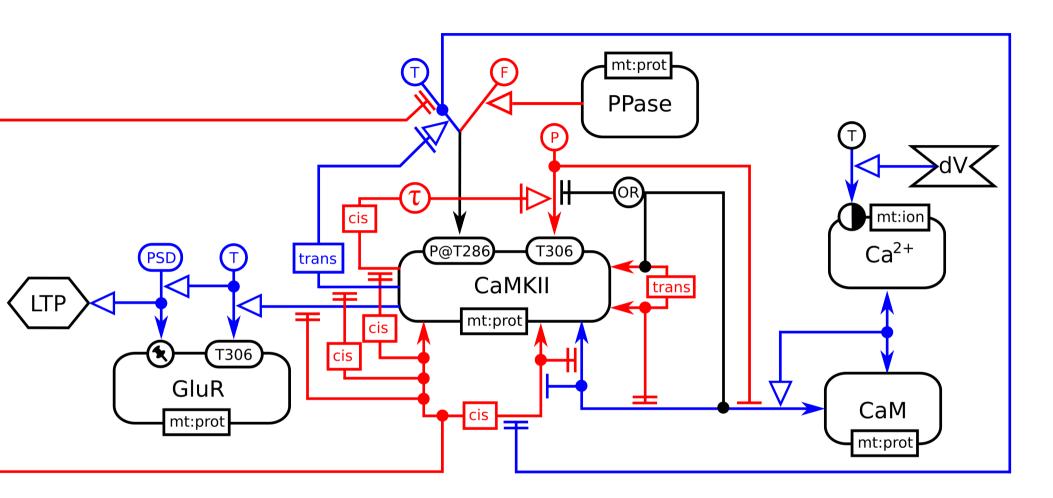
A necessary stimulation (Section 2.4.3.4) linking an entity node E and a relationship R means: "R only takes place if E exists.

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

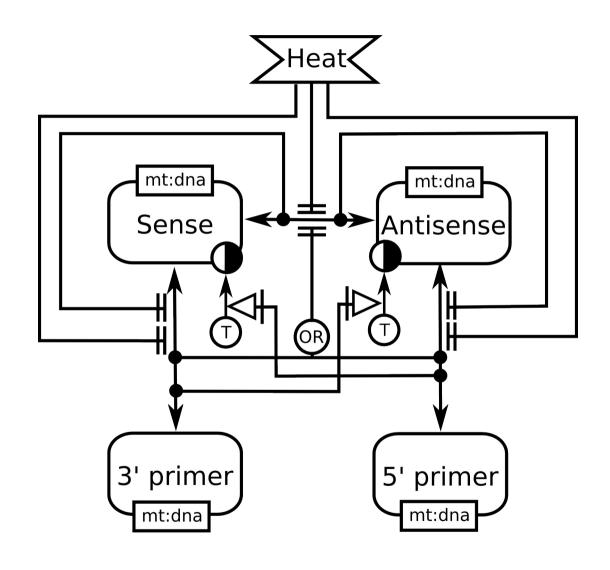


increases synaptic weight

decreases synaptic weight



ER map representing PCR



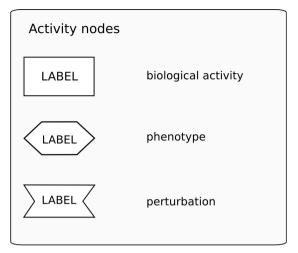


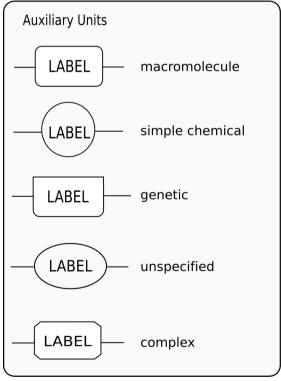
Pending issues

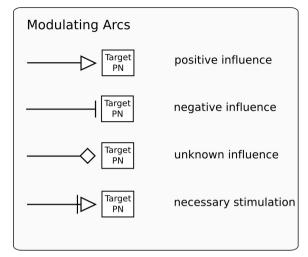
- Internal structure of entities, such as domains, sites, features. (NB: This has been agreed upon for the next version. We will be able to "nest" entities)
- Identification of instances: How to differentiate between several instances of the same entity, differentially involved in a relationships (e.g. trans-phosphorylation)?
- Identification of generics: How to lump together several entities for a given relationships? (e.g. MAPK instead of ERK1 and ERK2).

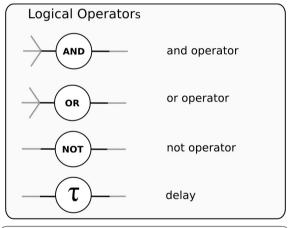


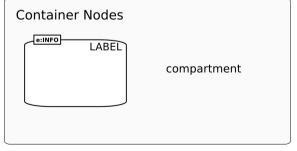
SBGN Activity Flows L1 reference card

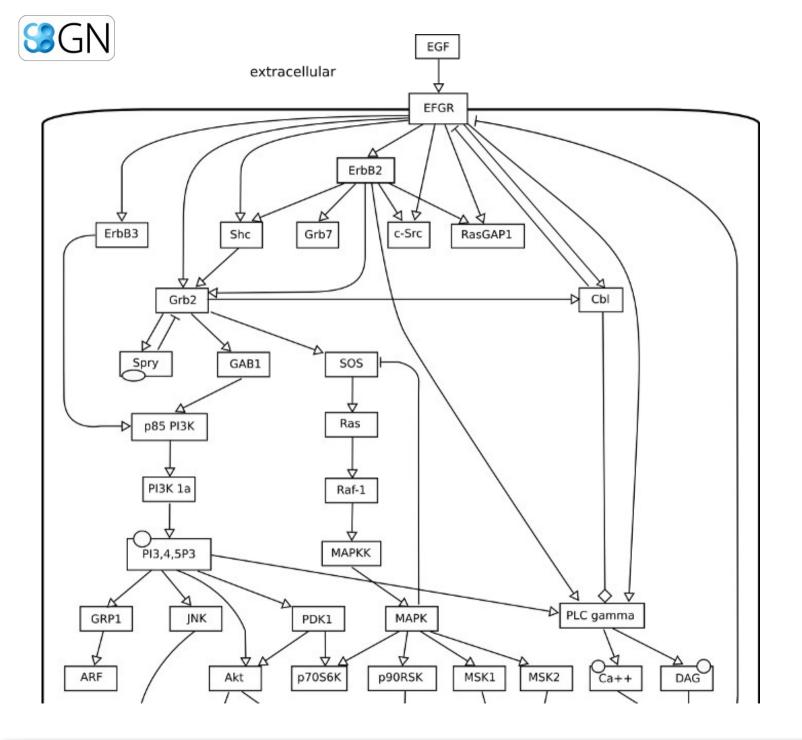










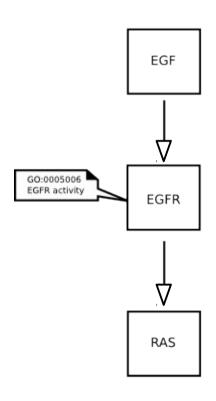


Example of Activity Flow map

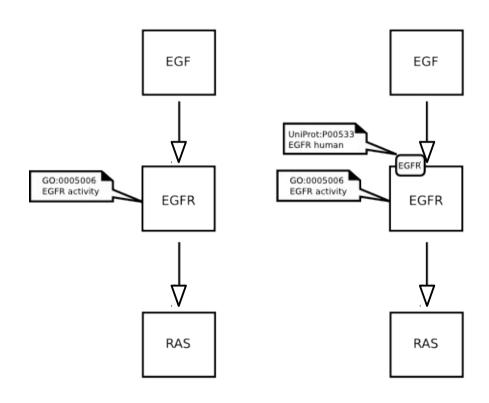


There Is More Than One Way To Do It (TIMTOWDI)

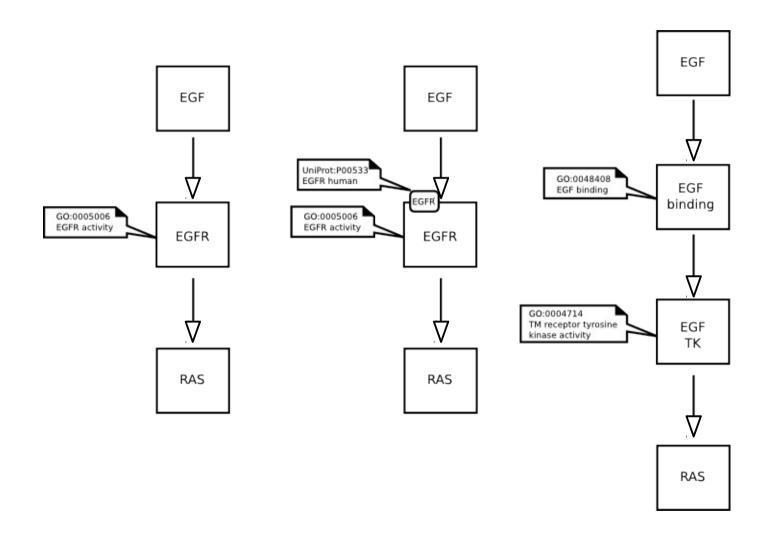




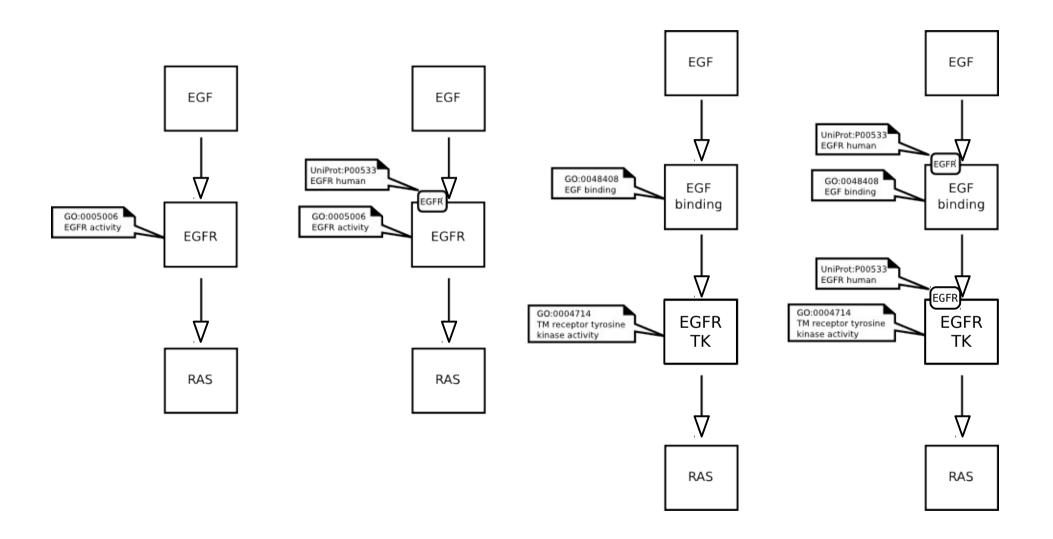


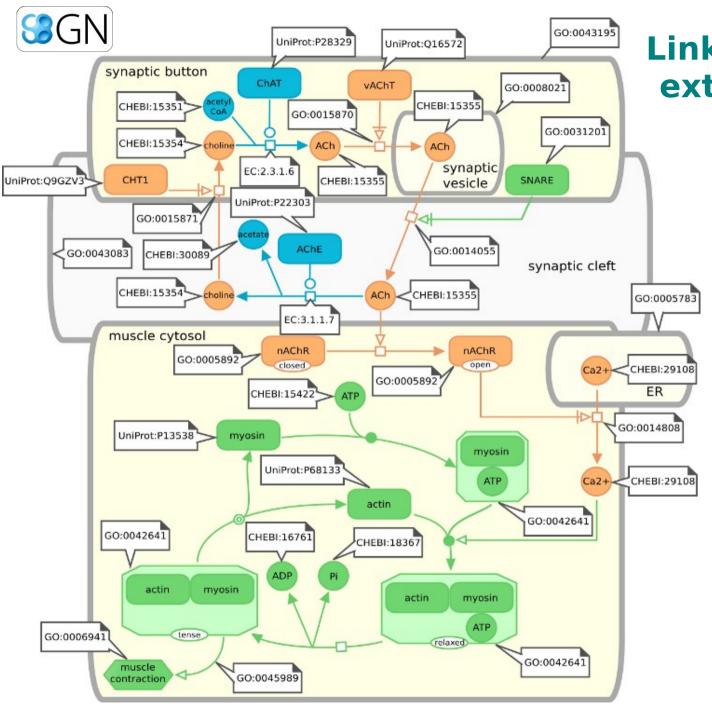




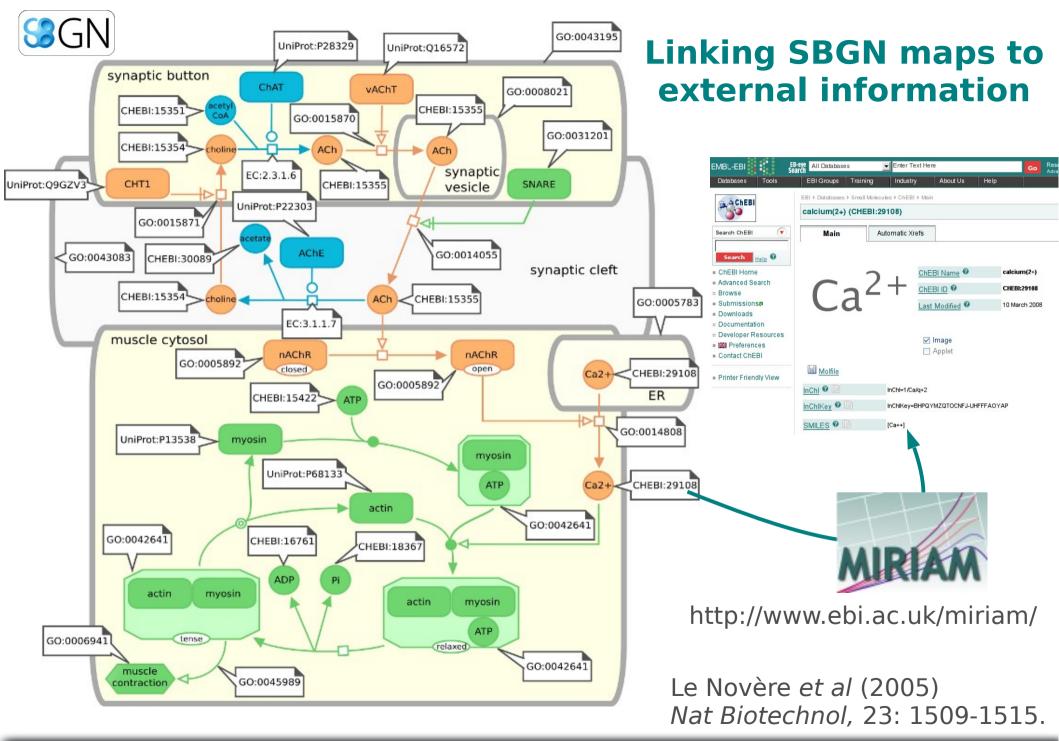


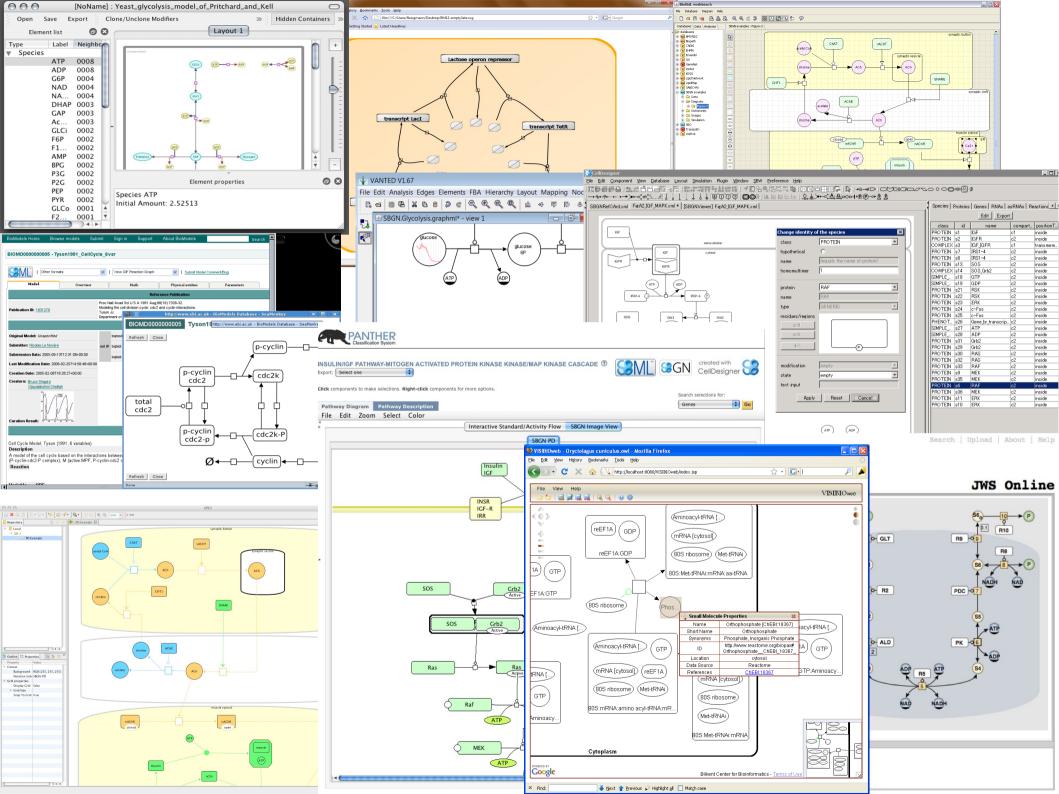






Linking SBGN maps to external information







Status of Specifications

- SBGN Process Descriptions
 - Level 1 Version 1.0 released on August 23rd 2008
 - Level 1 Version 1.1 released on September 1st 2009
 - Level 1 Version 2 under discussion
- SBGN Entity Relationships
 - Level 1 Version 1.0 released on September 1st 2009
- SBGN Activity Flows
 - Level 1 Version 1.0 released on September 1st 2009



Future SBGN meetings

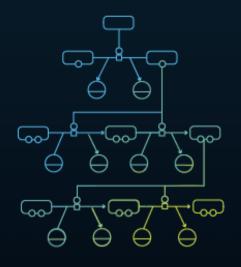
- 4rd SBGN hackathon (SBGN 5.5)
 - 21-23 April 2010
 - Wittenberg
- 6th SBGN forum (provisional)
 - October 2010
 - Edinburgh
 - Satellite of ICSB 2010



- Future hackathons and forums
 - part of HARMONY and COMBINE meetings
 - SBML (model), SED-ML (simulation), SBO (semantics), MIRIAM (metadata), SBGN (graphical), BioModels DB (distribution), ?



Home News Documents Lists Community Events About Q Google Site Search.



A Visual Notation for Network Diagrams in Biology

SBGN.org is the global portal for documentation, news, and other information about the Systems Biology Graphical Notation (SBGN) project, an effort to standardize the graphical notation used in maps of biochemical and cellular processes studied in systems biology.

Standardizing the visual representation is crucial for more efficient and accurate transmission of biological knowledge between different communities in research, education, publishing, and more. When biologists are as familiar with the notation as electronics engineers are familiar with the notation of circuit schematics, they can save the time and effort required to familiarize themselves with different notations, and instead spend more time thinking about the biology being depicted.

SBGN is made up of [] three orthogonal languages], representing different visions of biological systems. Each language defines a comprehensive set of symbols with precise semantics, together with detailed syntactic rules how maps are to be interpreted.

On this site, you can browse some <u>example maps</u> to get a feeling for SBGN, read the SBGN <u>specification documents</u>, join <u>online discussions</u>, see current working documents and software support in the SBGN wiki, and much more.

SBGN is the work of many people. It would not have been possible without the generous <u>support of multiple organizations</u> over the years, for which we are very thankful.

To quote SBGN as a whole, please use:

Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villéger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H. The Systems Biology Graphical Notation. Nat Biotechnol. 2009 27(8):735-41.

SBGN News

(02 Sep. '09) The first specifications for <u>SBGN Entity</u> <u>Relationships</u> and <u>SBGN</u> <u>Activity Flows</u>] are out.



What happens if one can read the blueprint









Acknowledgements

SBGN editors

- Nicolas Le Novère (UK)
- Stuart Moodie (UK)
- Anatoly Sorokin (UK)
- Michael Hucka (US)
- Falk Schreiber (DE)
- Huaiyu Mi (US)

Special contributors

Emek Demir, Katja Wegner, Mirit Aladjem, Sarala Wimalaratne, Frank Bergman, Ralph Gauges, Peter Ghazal, Kawaji Hideya, Lu Li, Yukiko Matsuoka, Alice Villéger, Sarah Boyd, Laurence Calzone, Melanie Courtot, Ugur Dogrusoz, Akira Funahashi, Sohyoung Kim, Fedor Kolpakov, Augustin Luna, Sven Sahle, Douglas Kell, Kurt Kohn, Hiroaki Kitano

The whole community participating to SBGN meetings, and mailing-lists







