

Systems biology

sbgn—a TikZ library to draw SBGN maps

Adrien Rougny  ^{1,2,*}

¹Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST) and ²Com. Bio Big Data Open Innovation Lab. (CBBDOIL), AIST, Aomi, Tokyo 135-0064, Japan

*To whom correspondence should be addressed.

Associate Editor: Alfonso Valencia

Received on June 7, 2018; revised on March 13, 2019; editorial decision on April 13, 2019; accepted on April 17, 2019

Abstract

Summary: The systems biology graphical notation (SBGN) has emerged as the main standard to represent biological maps graphically. It comprises three complementary languages: Process Description, for detailed biomolecular processes; Activity Flow, for influences of biological activities and Entity Relationship, for independent relations shared among biological entities. On the other hand, TikZ is one of the most commonly used package to ‘program’ graphics within TEX/LATEX. Here, we present *sbgn*, a TikZ library that allows drawing and customizing SBGN maps directly into TEX/LATEX documents, using the TikZ language. *sbgn* supports all glyphs of the three SBGN languages, and offers options that facilitate the drawing of complex glyph assembly within TikZ. Furthermore, *sbgn* is provided together with a converter that allows transforming any SBGN map stored under the SBGN Markup Language into a TikZ picture, or rendering it directly into a PDF file.

Availability and implementation: *sbgn*, the SBGN-ML to *sbgn* converter, as well as a complete documentation can be freely downloaded from <https://github.com/Adrienrougny/sbgn>. The library and the converter are compatible with all recent operating systems, including Windows, MacOS, and all common Linux distributions.

Contact: rougny.adrien@aist.go.jp

Supplementary information: [Supplementary material](#) is available at *Bioinformatics* online.

1 Introduction

Representing knowledge under the form of drawn diagrams is of crucial importance in Systems Biology, as it helps unveiling a holistic view on the biomolecular mechanisms underlying the biological process under study. Amongst different graphical representations, the systems biology graphical notation (SBGN) (Le Novère *et al.*, 2009) has emerged as the main standard to draw biological maps. It includes three orthogonal languages: process description (PD), activity flow (AF) and entity relationship (ER). PD allows representing precise biomolecular mechanisms under the form of processes; AF influences between biological activities and ER relationships between features of biological entities. SBGN maps can be visualized and edited by a growing number of software (see <http://sbgn.org> for a full list), and easily exchanged using the XML-based SBGN markup language format (SBGN-ML) (Van Iersel *et al.*, 2012). On the other hand, TikZ is one of the most commonly used package to ‘program’ graphics into TEX/LATEX

documents. It uses a straightforward syntax that is most convenient to draw graphs.

Here, we present *sbgn*, a TikZ library that allows drawing SBGN maps directly into TEX/LATEX documents using TikZ’s syntax. All three SBGN languages are supported, and all features offered by TikZ (such as setting the size of nodes or the style of nodes and arcs) can be used to draw highly customized maps. Additionally, *sbgn* is provided with an SBGN-ML converter that allows encoding a map stored under the SBGN-ML format into TikZ code, or rendering it directly under the form of a portable document format (PDF) image.

2 Drawing SBGN maps using *sbgn*

Formally, SBGN maps can be seen as graphs: Each of the three SBGN languages defines a set of glyphs (i.e. graphical symbols) representing either nodes (biological entities and processes), arcs

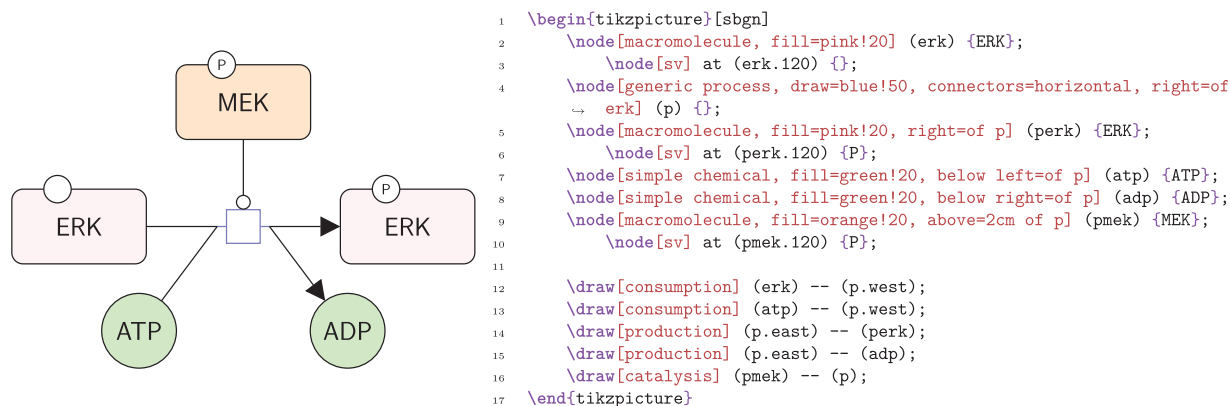


Fig. 1. Drawing a map using *sbgntikz*. Right panel shows a sample of TikZ code using *sbgntikz*. This sample encodes the PD map rendered in the left panel. SBGN nodes and decorators are drawn using the TikZ's `\node` command (lines 2–10), while arcs are drawn using the `\draw` command (lines 12–16). A glyph is drawn using its associated key and other optional keys that modify its appearance. Glyphs can be positioned relatively to each other using TikZ's positioning library (lines 3–10)

(relationships) or decorators (attributes). Figure 1 gives an example of PD map (left panel). This map represents the phosphorylation of ERK, catalyzed by phosphorylated MEK. Each type of glyph represents a particular biological concept. For example, in the above map, rectangular nodes with rounded corners represent pools of macromolecules, whereas the arc with a circle-shaped arrowhead represents a catalysis, and the circles decorating rectangles represent state variables.

sbgntikz implements all SBGN glyphs in TikZ by defining new TikZ styles and portable graphics format (PGF) shapes, and associates with each glyph a custom TikZ key that can be called in a TikZ picture in order to draw that glyph (e.g. `macromolecule` for the macromolecule glyph, `sv` for the state variable glyph). Figure 1 gives a sample of TikZ code using *sbgntikz* (right panel), and the rendered map obtained when compiling this code (left panel). SBGN nodes and decorators are straightforwardly drawn using TikZ's `\node` command (lines 2–10), and arcs using the `\draw` command (lines 12–16). *sbgntikz* defines additional keys that allow modifying the way glyphs are drawn. For example, the `connectors` key, when supplied to the `\node` command together with the `generic process` key, makes the TEX engine draw vertical or horizontal connectors that stick out of each side of a process node (line 4).

All features offered by default in TikZ and by additional TikZ libraries (such as the positioning and fit libraries) are compatible with *sbgntikz*. Therefore, SBGN maps drawn using *sbgntikz* can be customized at will. For example, the size and draw/fill color of glyphs, the font of labels, etc. can be changed either locally, for a certain type of glyph, or globally, by using the appropriate TikZ keys; or glyphs can be placed relatively to each other on the map (lines 3–10).

3 SBGN-ML to TikZ converter

The SBGN-ML is an XML-based format that has been developed to store and exchange SBGN maps. A number of popular SBGN visualization and edition software support SBGN-ML for reading and/or writing maps [SBGN-ED (Czauderna *et al.*, 2010), Newt (Sari *et al.*, 2015), etc.], and some pathway databases offer an SBGN-ML export option [Reactome (Croft *et al.*, 2013), Path2Models (Büchel *et al.*, 2013), etc.]. In order to be able to integrate SBGN-ML maps into TEX/LATEX files and take advantage of TikZ's customization features, we implemented an SBGN-ML to TikZ converter in Python. This converter allows encoding an SBGN-ML map into TikZ code

using *sbgntikz*, and optionally rendering it under the form of a PDF image. It supports all features of SBGN-ML, and uses some relative positioning to ease subsequent edition of output *sbgntikz* maps. The converter can be called directly from the command line to output a TEX file or a PDF image, or integrated into other software as a library. Multiple examples of SBGN maps rendered using the converter are available in the [Supplementary Material](#).

4 Software implementation and availability

sbgntikz is implemented in TEX/LATEX using the PGF/TikZ package, and the converter is implemented in Python 3. *sbgntikz*, the converter and a complete documentation are freely available from <https://github.com/Adrienrouigny/sbgntikz>. Both the library and the converter can be modified and redistributed under the terms of the GPLv3 license.

Acknowledgements

The author would like to thank the SBGN community for the development of SBGN and related software, and Till Tantau for the development of the PGF/TikZ package and its extraordinarily rich documentation.

Funding

This work was supported by the Com. Bio Big Data Open Innovation Lab. (CBBD-OIL).

Conflict of Interest: none declared.

References

- Büchel, F. *et al.* (2013) Path2Models: large-scale generation of computational models from biochemical pathway maps. *BMC Syst. Biol.*, **7**, 1.
- Croft, D. *et al.* (2013) The reactome pathway knowledgebase. *Nucleic Acids Res.*, **42**, D472–D477.
- Czauderna, T. *et al.* (2010) Editing, validating and translating of SBGN maps. *Bioinformatics*, **26**, 2340–2341.
- Le Novère, N. *et al.* (2009) The systems biology graphical notation. *Nat. Biotechnol.*, **27**, 735–741.
- Sari, M. *et al.* (2015) SBGNviz: a tool for visualization and complexity management of SBGN process description maps. *PLoS One*, **10**, e0128985.
- Van Iersel, M.P. *et al.* (2012) Software support for SBGN maps: SBGN-ML and LibSBGN. *Bioinformatics*, **28**, 2016–2021.