the combine
computational modeling in biology network

HARMONY 2024

ABSTRACT BOOKLET

Note: Lightning talk/intro session on Monday 8 Apr can be joined remotely here.

Lightning talks

Presenting Author: Manuel Lera-Ramirez, University College London

Authors: Manuel Lera-Ramirez Bähler, Jürg (University College London)

ShareYourCloning: towards a FAIR standard for sequence provenance

Abstract:
Recombinant DNA technology is used in diverse research fields to generate new DNA sequences by combining fragments from existing DNA molecules. These engineered molecules are introduced into cells as self-replicating plasmids and can be used to modify their genomes.
Currently, no established standard exists to describe how DNA sequences are generated from its ancestors. Therefore, researchers document the provenance of plasmids, cell lines and strains using spreadsheets or text-based systems, which are necessarily inconsistent and differ between collections. Some proprietary tools keep track of the provenance of sequences, but they don't allow users to export this information in an open format.
ShareYourCloning aims to develop a data model to describe sequence provenance while developing web tools that leverage the model and allow researchers to plan and document their experiments. In this way, the data is captured through an intuitive interface in a standardised interoperable format at the beginning of the data lifecycle. For now the data is exported as json, but in the near future it will be possible to export a sequence and its provenance in SBOL format, following SBOL's PROV-O guidelines.
You can try the web application at https://shareyourcloning.netlify.app

---

Presenting Author: Paola Di Maio, IGDORE

Authors: Paola Di Maio

Neuroscience Model Card ML (NMCML

Abstract:
There are many approaches to modelling the Brain. NeuroML does a great job at capturing some of them. Neuroscience datasets however are expanding rapidly, and the diversity of facets, models and perspectives to modelling the brain it is very complex. It can be disorienting and mindwrecking even for experienced bioinformaticians. One approach to handle this complexity is the use of Model Cards
{https://doi.org/10.1007/978-3-030-86993-9_8) In this poster and lightning talk I present my work in using Model Cards and in the formulation of schemas, based on the newest and largest brain dataset, since I put my hands on a human readable version courtesy of Allen Institute. Currently on early versions of model cards and seeking feedback on beta versions
{https://huggingface.co/STARBORN/modelcardML_V1)

Presenting Author: V Lokesh, Indian Institute of Technology Bombay

Authors: V Lokesh Wangikar, Pramod

Metabolic modeling of novel fast growing cyanobacterium Synechococcus elongatus PCC 11801 for value added products

Abstract:
Greenhouse gases from chemical manufacturing industries contribute significantly to climate change and global warming, and their sequestration is imperative to reducing the carbon footprint. Cyanobacteria are photoautotrophic microorganisms capable of fixing CO2 from the environment via photosynthesis, and can be engineered to convert the CO2 to value-added products. Despite progress in synthetic biology and strain engineering technologies, the commercial applicability of cyanobacteria has been limited due to the slow growth rate. However, the freshwater strain Synechococcus elongatus PCC 11801 recently isolated from our lab is one of the fastest growing strain with tolerance to different environmental stresses, making it a potential cell factory. To streamline and guide strain engineering efforts for PCC 11801, we generated a genome scale metabolic model for the strain that can be utilized with flux analysis to generate in silico predictions of hotspots for genetic perturbation that would maximize a specific value-added product in its metabolic network. The model was reconstructed using the standard protocol of draft reconstruction generation, manual curation, model debugging and validation. The genome scale model (‘iJB785’) of a phylogenetically close model strain Synechococcus elongatus PCC 7942 was used as the starting point. The draft reconstruction of PCC 11801 was generated by first mapping the homologous genes with iJB785 through bidirectional best hit analysis of the genomes. The model reactions were manually curated to reduce redundancy, and verified with the BiGG, KEGG and BRENDA database. 287 novel reactions identified pertinent to PCC 11801 were added de novo into the model with the corresponding gene-protein-reaction association. Model debugging was performed through flux balance analysis, flux variability analysis, single gene deletion analysis, and single reaction deletion analysis. Overall, the metabolic model consisted of 1052 genes and 1130 reactions. The model was validated by gene essentiality analysis with a sensitivity and specificity score of 0.9223 and 0.5425 respectively. The metabolic robustness of the model was estimated to be 39% from the model, which validates the amenability of the strain for genetic engineering. The model predicted that the strain would convert CO2 to succinic acid at a higher production rate compared to other value-added products through the in silico predictions.

Presenting Author: Francis Chemorion, University of Pompeu Fabra

Authors: Francis Chemorion

Enhancing Computational Biology Reproducibility and Accessibility through a Docker-Based Registry for Biomodels

Abstract:
The reproducibility of computational models in biology is hindered by the variability in execution environments, leading to challenges in model sharing and reuse. This work introduces a novel Docker-based registry designed to distribute Systems Biology Markup Language (SBML) models and Simulation Experiment Description Markup Language (SED-ML) files, along with their auto-generated interfaces and solvers. By encapsulating these elements in Docker containers, we ensure consistent execution across different computing environments, thus addressing a significant barrier to reproducibility and accessibility. Our approach not only facilitates the seamless sharing of computational models but also promotes the adoption of standardized practices within the scientific community. An evaluation of the registry underscores its potential to improve usability, foster collaboration, and enhance the integrity of computational research in biology. This initiative aligns with efforts to standardize computational models and their dissemination, promising to advance the field of computational biology by making robust, reproducible research more attainable.

Presenting Author: Goksel Misirli, Keele University

Authors: Goksel Misirli

libSBOLj3: a graph-based library for design and data exchange in synthetic biology

Abstract:
The Synthetic Biology Open Language version 3 data standard uses a graph-based approach to exchange biological design information, offering novel features for the reuse of information and an extended terminology for biological constraints. Here, we present libSBOLj3, a Java library to facilitate the development of computer-aided design and automation tools, data exchange and interoperability between different software tools. The library can be integrated into other software tools and computational workflows. The application programming interface is built on graph concepts for information search and retrieval, and it can incorporate application-specific data. Furthermore, this library provides a validation layer to ensure the creation of compliant documents by validating against required and conformance-related rules and best practices.
Paper: https://doi.org/10.1093/bioinformatics/btad525
Availability: https://github.com/SynBioDex/libSBOLj3

Presenting Author: Frank T. Bergmann, BioQUANT, Heidelberg University

Authors: Frank T. Bergmann

COPASI.js: How to easily use COPASI from the web

Abstract:
COPASI.js is a JavaScript library that enables seamless integration of COPASI, a modeling and simulation platform utilized in systems biology, biochemistry, and related fields with the web. With COPASI.js, users can harness the power of COPASI directly within web-based applications, allowing for dynamic and interactive modeling, simulation, and analysis tasks. https://copasijs.readthedocs.io/en/latest/

Presenting Author: T.J. Sego, University of Florida

Authors: T.J. Sego Rahuman Sheriff

Towards Reproducible Stochastic Biological Simulation

Abstract:
Stochastic simulations are commonly used to quantitatively or semi-quantitatively describe the dynamics of biological systems. At various scales and in multiple applications, stochastic simulation better reflects observed biological processes and robustness. Various methods are widely used to incorporate stochasticity into biological simulation, such as the Gillespie stochastic simulation algorithm for systems biology modeling, stochastic Boolean networks for network modeling, and the Cellular Potts model methodology for multicellular modeling. Proving reproducibility of simulation results is critical to establishing the credibility of a model. To this end, BioModels, the largest repository of curated mathematical models, tests and reports the reproducibility of simulation results for all submitted models when possible. A recent study showed that about 50% of the deterministic ordinary differential equation models on BioModels could not be reproduced when applying criteria for reproducibility to the information provided in their associated publication, reflecting a current crisis of reproducibility. Furthermore, there are no well-accepted metrics or standards for reproducing stochastic simulation results, thus perpetuating the crisis of reproducibility for a broad class of biological models. This lightning talk survey recent progress to establish an accepted framework for testing the reproducibility of stochastic simulations in biological modeling. The talk will provide an overview of recent progress towards defining quantitative measures to determine whether stochastic simulation results can be reproduced, and when results have been reproduced. The talk complements a breakout session at HARMONY 2024 that intends to continue the described work and invite new collaborators to join continuing efforts.

Presenting Author: Carolus Vitalis, University of Colorado Boulder

Authors: Carolus Vitalis Amos, Tyler (University of Colorado Boulder); Hanel, Juan (University of Colorado
Boulder); Acuña, Daniel (University of Colorado Boulder); Myers, Chris (University of Colorado Boulder)

PUMA: Promoter Unraveling through Machine-learning Algorithms

Abstract:
Synthetic Biology aims to apply the engineering principles of standards, abstraction, and decoupling to genetic engineering. In the interest of the first, synthetic biologists have developed notations and standardized DNA sequences for the different genetic parts, like promoters, ribosome binding sites, and coding sequences, among others, which are then collected and deposited in repositories like the iGEM Part Repository or SynBioHub. While these repositories offer a lot of different genetic parts and could serve as a good starting point; they tend to lack curation of the parts deposited, which makes the identification, classification, and finding of the ideal parts a hassle.
To overcome this challenge, we aim to develop an AI model capable of recognizing a DNA sequence corresponding to a Promoter part from the iGEM Repository. This could help make Synthetic Biology a more accessible field, greatly accelerating the rate of research and novel applications in a wide array of fields, from healthcare to agriculture and even space colonization.
The model would be trained using the Promoter sequences available in the 2019 iGEM Part Distribution, which contains the DNA sequence of 230 hand-curated promoters, among other parts. After the model has been trained, we expect to test it on the full iGEM Repository to ensure that the model can identify valid promoters, even if they are part of a longer DNA sequence.

Presenting Author: Xiaoming Hu, Heidelberg Institute for Theoretical Studies

Authors: Xiaoming Hu Golebiewski, Martin (Heidelberg Institute for Theoretical Studies, Heidelberg, Germany)
Meineke, Frank (Leipzig University Medical Center, Leipzig, Germany)
Abaza, Haitham (Heidelberg Institute for Theoretical Studies, Heidelberg, Germany)
Löbe, Matthias (IMISE University of Leipzig, Leipzig, Germany)
Haensel, Rene (IMISE University of Leipzig, Leipzig, Germany)
Wu, Qi (Heidelberg Institute for Theoretical Studies, Heidelberg, Germany)
Müller, Wolfgang (Heidelberg Institute for Theoretical Studies, Heidelberg, Germany)

NFDI4Health Local Data Hubs for Finding and Accessing Personal Health Data

Abstract:
NFDI4Health, a vital component of Germany's NFDI network, receives joint funding from federal and state governments through the German Research Foundation. Dedicated to fostering ethical and private data sharing in clinical and epidemiological research, NFDI4Health aligns with FAIR principles. Its metadata schema facilitates standardized publication of health study metadata across services like the German Central Health Study Hub (GCHSH) and Local Data Hubs (LDH).
The Local Data Hub (LDH) serves as the primary local connector for NFDI4Health's federated distributed services. Operating within and beyond Data Holding Organizations (DHOs), the LDH enables local data structuring, bundling, and sharing. It operationalizes NFDI4Health standards, employing a modularly designed metadata schema (MDS) tailored for health studies, facilitating seamless communication with the GCHSH. Integration into DHO processes is imperative for streamlined workflows.
Chosen for its adherence to FAIR principles, FAIRDOM-SEEK stands out as the preferred LDH platform. Its well-established user community, compatibility with NFDI4Health's requirements, and demonstrated success in data management reinforce its suitability. The10 decision is further supported by HITS' development expertise and their pivotal role in co-founding the FAIRDOM initiative.
To address Metadata Schema (MDS) complexity, SEEK incorporates an ""extended metadata"" feature now implemented in the LDH. This feature allows customizable user-defined extensions, tailored to meet NFDI4Health's needs, ensuring a comprehensive representation of the metadata schema. Extended Metadata enhances the description of complex metadata structures within the FAIRDOM Community, strengthening the LDH's pivotal role in advancing FAIR principles in health research data.

Presenting Author: Natasa Miskov-Zivanov, University of Pittsburgh

Authors: Natasa Miskov-Zivanov Emilee Holtzapple, Casey Hansen

MINUET: Automated and systematic verification to increase quality and long-term reuse of models

Abstract:
Although modeling is an important component of a research pipeline in biology, most often there is no systematic or standardized approach for quality assessment and annotation of models, reducing their trustworthiness and reuse potential. Moreover, most of the model design and documenting steps are still done manually. Creating useful and reliable models of cellular signaling requires thorough and careful information extraction, knowledge assembly, and comprehensive model verification and validation. Manual execution of these steps can be highly impractical due to their complexity and the detailed attention they require. The verification step, assessing whether the model structure is correct by finding support for all its elements and interactions, and the validation step, evaluation of model behavior against experimental observations and data, usually occur iteratively with model expansion before the model can be used to make predictions or explanations.
In this work, we introduce MINUET (Managing Interaction and Network (re-)Usability through Evaluation of Trustworthiness), a platform for automated network verification and curation. MINUET is comprehensive, conducting model verification that relies on both knowledge and data, the vast published literature and publicly available data-bases; it is flexible as it can conduct both context-aware and context-independent verification; it is versatile, allowing users to conduct in-design verification during model creation, post-design verification of existing models, or a comparison of models to verify them against each other; finally, it is fast, due to its automated steps for retrieving and comparing interactions. MINUET contributes to `the four R's` in several ways: it evaluates and facilitates reliability and reusability of models and information; it strengthens the potential for reproducibility of predictions across models by identifying structural differences between them; and it assesses the replicability of outcomes and observations by collecting evidence from knowledge and data sources.

Presenting Author: Gaoxiang Zhou, University of Pittsburgh

Authors: Gaoxiang Zhou Sayed, Khaled (University of New Haven);
Luo, Haomiao (University of Pittsburgh);
Arazkhani, Niloofar (University of Pittsburgh);
Ahmed, Yasmine (University of Pittsburgh);
Holtzapple, Emilee (University of Pittsburgh);
Tang, Difei (University of Pittsburgh);

Bocan, Kara (University of Pittsburgh);
Telmer, Cheryl (Carnegie Mellon University);
Miskov-Zivanov, Natasa (University of Pittsburgh);

A flexible modeling and simulation framework for hybrid elements

Abstract:
To retain the benefits of both logical and discrete modeling, and to adapt to the varying precision or amount of the available information, we introduce here the DiSH2.0 (Discrete, Stochastic, Heterogeneous) modeling and simulation framework. In DiSH2.0, both logical and algebraic update functions can exist within a hybrid model, enabling optimal use of the retrieved element and interaction information, including qualitative and causal information, as well as quantitative information when available. Instead of encoding discrete element values with Boolean variables, using discrete variables in DiSH2.0 allows elements within a single model to have any necessary number of discrete levels and simplifies model creation, verification, and use. Additionally, the models have `memory`, as element update rules include the previous value of an element when calculating its change. The use of regulation weights and other notations in update rules offers a way to account for varying effects across different regulators as well as different levels of a single regulator.
For these new hybrid models, we introduced spontaneous increase and decrease behaviors, different responses to the balance between negative and positive regulators, and varying increment/decrement amounts. These features offer the flexibility to increase model accuracy when precise information is available or to retain a more abstract level in the absence of such information. We implemented a simulator that supports these new features and is also compatible with our previous work and other discrete modeling tools. The simulator expands to hybrid models the suite of deterministic and stochastic (also known as synchronous and random asynchronous, respectively) simulation schemes that have been mainly used for logical models in the past.
We demonstrated on a T cell differentiation use case the accuracy and precision improvements provided by hybrid modeling and new modeling and simulation features. Using the BioRECIPE representation format, which is both human and machine readable and editable, DiSH2.0 is compatible with other tools enabling full automation of the modeling pipeline: interaction extraction from literature and databases, interaction filtering, assembly of new and extension of existing interactions, and model verification and validation.
The repository of DiSH 2.0 is hosted at https://github.com/pitt-miskov-zivanov-lab/DiSH and its documentation at https://melody-dish.readthedocs.io

---

Presenting Author: Rob Vickerstaff, UCL

Authors: Rob Vickerstaff

Analysing BioModels Database using PyNeuroML

Abstract:

We present progress with automated model validation and testing of the EBI BioModels database using newly added features of the PyNeuroML tool. Model SBML and SedML files are downloaded, validated and run using the Tellurium simulator. Results are presented as a Markdown table with summary information.

Presenting Author: Michael Blinov, UConn Health

Authors: Michael Blinov

Visualization of Rule-Based Models

Abstract:
Rule-based modeling is a powerful method to describe and simulate interactions among multi-site molecules and multi-molecular species, accounting for the internal connectivity of molecules in chemical species. Numerous rule-based models are published annually, and the accompanying code (BioNetGen Language, BNGL) for these models is typically included in supplementary materials. However, BNGL is not fully human-readable, and models encoded in BNGL are difficult to comprehend without a steep learning curve. Here, we introduce bnglViz, an online platform for visualizing BNGL files as graphical cartoons, empowering researchers to grasp the nuances of rule-based models swiftly and efficiently and making the exploration of complex biological systems more accessible than ever before. bnglViz can be used to visualize published models and explore newly designed models created in non-graphical modeling tools
Web: http://MolClustPy.github.io
http://bnglViz.github.io
http://ModelBricks.org

Presenting Author: Matthias König, Humboldt-University Berlin, Faculty of Life Science, Institute for Biology, Systems Medicine of the Liver, https://livermetabolism.com

Authors: Matthias König

Enhancing Computational Model Development in Systems Biology Using SBML: sbmlutils, sbm4humans, cy3sbml

Abstract:
The Systems Biology Markup Language (SBML) [1] is recognized as the standard framework for representing and exchanging complex mathematical models in biological systems research. SBML facilitates the depiction of a diverse array of biological phenomena, encompassing metabolic networks, signaling pathways, and regulatory networks. It is

versatile enough to handle models ranging from simple individual processes to intricate multi-scale representations.
One of the primary challenges faced by newcomers in computational biology is the encoding and development of ordinary differential equation (ODE) models within the SBML framework. Addressing this hurdle, we introduce two innovative Python tools: sbmlutils [2], sbml4humans [3], and the Cytoscape application cy3sbml [4]. These tools collectively streamline the process of SBML model creation, enhancing both the programmatic aspect and the user experience. Specifically, sbmlutils facilitates the programmatic construction of SBML models, while sbml4humans generates user-friendly reports for model interpretation. Furthermore, cy3sbml integrates with Cytoscape to offer advanced visualization capabilities, thereby augmenting the comprehension and analysis of SBML-encoded models.
These advancements significantly contribute to the ease of SBML model development and interpretation, fostering greater accessibility and understanding for those entering the field of computational systems biology.
[1] Keating SM, Waltemath D, König M, Zhang F, Dräger A, Chaouiya C, Bergmann FT, Finney A, Gillespie CS, Helikar T, Hoops S, Malik-Sheriff RS, Moodie SL, Moraru II, Myers CJ, Naldi A, Olivier BG, Sahle S, Schaff JC, Smith LP, Swat MJ, Thieffry D, Watanabe L, Wilkinson DJ, Blinov ML, Begley K, Faeder JR, Gómez HF, Hamm TM, Inagaki Y, Liebermeister W, Lister AL, Lucio D, Mjolsness E, Proctor CJ, Raman K, Rodriguez N, Shaffer CA, Shapiro BE, Stelling J, Swainston N, Tanimura N, Wagner J, Meier-Schellersheim M, Sauro HM, Palsson B, Bolouri H, Kitano H, Funahashi A, Hermjakob H, Doyle JC, Hucka M; SBML Level 3 Community members. SBML Level 3: an extensible format for the exchange and reuse of biological models. Mol Syst Biol. 2020 Aug;16(8):e9110. doi: 10.15252/msb.20199110.
[2] sbmlutils - Python utilities for SBML, https://github.com/matthiaskoenig/sbmlutils
[3] sbm4human - Human readable SBML reports, https://sbml4humans.de
[4] cy3sbml - Visualization of SBML models in Cytoscape, https://github.com/matthiaskoenig/cy3sbml

---

Presenting Author: Matthias König, Humboldt-University Berlin, Faculty of Life Science, Institute for Biology, Systems Medicine of the Liver, https://livermetabolism.com

Authors: Matthias König

Advancing Liver Function Assessment: Personalized and Stratified Approaches with Standardized Computational Models and Data

Abstract:
Essential prerequisites for the practical application and translation of computational models include: i) reproducibility of results; ii) reusability and extensibility of models; iii) data availability; and iv) strategies for model stratification and individualization. In this study, we present a modeling workflow tailored to these critical aspects, with a focus on liver function tests.
Evaluating liver function is a crucial task in hepatology, yet accurately quantifying hepatic function has persisted as a clinical challenge. Dynamic liver function tests offer a promising

method for non-invasive in vivo assessment of liver function and metabolic phenotyping. These clinical tests determine liver function through the elimination of a specific test substance, thus revealing information about the liver's metabolic capacity.
We employed whole-body physiologically-based pharmacokinetic (PBPK) models to simulate these tests, which encompass absorption, distribution, metabolism, and elimination processes. PBPK models serve as powerful instruments for investigating drug metabolism and its impact on the human body. Here we will provide a short overview of our efforts in utilizing PBPK models as digital twins for metabolic phenotyping and liver function evaluation. To develop and validate our models, we created the first open pharmacokinetics database, PK-DB, containing curated data from over 600 clinical studies. Our models are individualizable and stratifiable, enabling simulation of lifestyle factors and co-administration effects on drug metabolism.
We have applied our models to various clinical inquiries, such as simulating individual outcomes post-hepatectomy using an indocyanine green model and examining the influence of CYP2D6 gene variants through a dextromethorphan model integrated with drug-gene interactions. These models are constructed hierarchically, describing metabolic and other biological processes in organs like the liver and kidneys, connected to whole-body physiology. All models and data are accessible for reuse in a reproducible manner, encoded in the Systems Biology Markup Language (SBML).
In this study, we provide an overview of PBPK models and demonstrate how SBML, COMBINE standards, and FAIR principles can facilitate model development, coupling, and reuse.

Presenting Author: Eva Liu, University of Washington

Authors: Eva Liu Sai Anish Konanki, Edison Shunming Eshao, Sam Chou, Lucian P. Smith, Herbert M Sauro, Joseph L Hellerstein

An Intelligent Antimony Web Editor: A Zero-Install Source Editor for Antimony Models

Abstract:
Developing biochemical models in systems biology is a complex, knowledge-intensive activity. Some modelers (especially novices) benefit from model development tools with a graphical user interface. However, as with the development of complex software, text-based representations of models provide many benefits for advanced model development.
We have previously developed and released the VSCode-Antimony editor, a tool for building, analyzing, and translating models written in the Antimony modeling language, a human readable representation of Systems Biology Markup Language (SBML) models. VSCode-Antimony is a source editor, a tool with language-aware features. For example, there is autocompletion of variable names to assist with model building, hover messages that aid in model analysis, and translation between XML and Antimony representations of SBML models. These features result from making VSCode-Antimony model-aware by incorporating several sophisticated capabilities: analysis of the Antimony grammar (e.g. to identify model symbols and their types); a query system for accessing

knowledge sources for chemical species and reactions; and automatic conversion between different model representations (e.g. between Antimony and SBML).
We have found that the reliance on VSCode has significant drawbacks in that users must install VSCode prior to installing the extensions for Antimony. A subtlety here is that our extension is developed in python, and so great care is required so that our extension does not conflict with python packages on the user's machine. These issues have led to our development of AWE, an Antimony Web Editor that is entirely browser-based and only requires a source of static pages (e.g., GitHub). The new version also incorporates the latest metadata extensions added to Antimony. As such, the user need only browse to the AWE URL to use the editor. This lightning talk discusses our beta of AWE.

Presenting Authors: Kerstin Gierend, Department of Biomedical Informatics, Medical Faculty Mannheim, Heidelberg University; Sascha Genehr, Institute of Communications Engineering, University of Rostock

MIRAPIE community project- a minimal biomedical data provenance model

Abstract:

Biomedical data is diverse and often highly complex. Hospitals, research institutions, and companies hold a multitude of data items gathered and processed differently. Many countries already defined standards for terminologies, formats, and common IT infrastructure for data exchange. However, even on national levels, the standardized format of the data does not necessarily result in comparable or integrable data content. Bias can be introduced during data collection and processing, emphasizing the critical need for a robust data provenance model to ensure the reliability and, in the best-case, reproducibility of study results, while also establishing trust. Provenance information provides a necessary basis for interpretation and usability evaluation of biomedical data elements and results for their secondary use. Minimizing the risk for misinterpretation is specifically crucial in biomedicine, as research results may have an impact on patient treatment.
We launched an international open community project, where data is contributed by multiple institutions like medical data integration centers, or clinical researchers with the aim to establish MInimal Requirements for Automated Provenance Information Enrichment (MIRAPIE). The project idea and essential requirements were already shared in the Provenance Week Conference 2023, together with an international call for contribution. Currently, community members from different European countries share their ongoing work at the public repository (https://codeberg.org/MIRAPIE/MIRAPIE).
Our minimal data provenance model emphasizes open community engagement, active participation in the development and refinement of the biomedical provenance standard. The model addresses emerging challenges in data provenance, especially related to data privacy restriction due to sensitive patient data in research. This model answers seven general questions at each data processing step from the WHO did WHAT (and HOW), WHERE, with

WHICH tool, WHEN and WHY perspective. Furthermore, it maps the questions to related definitions, contained in the PROV ontology.
In summary, this community tailored biomedical data provenance model aims to enhance transparency, traceability, and accountability of biomedical data throughout the data lifecycle. In addition, this provenance information is expected to facilitate the sharing of insights, thus contributing to the FAIR principles.

Presenting Author: Fengkai Zhang, NIAID/NIH

Authors: Fengkai Zhang Meier-Schellersheim, Martin (NIAID/NIH)

Exploring Parameter Space of Rule-based Models: The Functions and Features of Simmune Modules

Abstract:
The approach of rule-based modeling empowers modelers to construct intricate molecular interaction networks through the definition of specific 'rules' governing molecular interactions and complex state transitions. These rules encapsulate detailed aspects such as sub-molecular domain interactions, as well as the phosphorylation and binding states of the molecules involved. By embracing this methodology, our team's software suite, Simmune, facilitates the simulation of cellular processes in either well-stirred or spatially resolved environments. During this meeting, we will host a tutorial session to demonstrate how Simmune can be utilized to explore the parameter space of rule-based models. In this talk, we will provide a comprehensive overview of the modules within Simmune and their respective roles in simulating rule-based models. The development of Simmune is supported by the intramural program of the NIAID, NIH.

Presenting Author: Ion Moraru, UConn Health

Authors: Ion Moraru Schaff, James; Patrie, Alex; Drescher, Logan; Smith, Lucian; Agmon, Eran

BioSimulations 2.0

Abstract:
To help investigators share and reuse simulations, we developed BioSimulations, a central repository for models, simulations, and visualizations of simulation results. Importantly, BioSimulations both helps authors quality control and share their projects and helps others modify and execute these projects and interactively visualize their results. Already, BioSimulations 1.0 includes over 1,000 projects. BioSimulations supports a broad range of biological systems and scales as it incorporates numerous modeling approaches, model languages, model repositories, simulation algorithms, and simulation tools. We achieved this

by refining SED-ML and expanding SBO and KiSAO; developing new formats for simulation results, logs of simulations, data visualizations of simulation results, and simulation tools; developing BioSimulators, a registry of simulation tools; developing tools for quality controlling simulations and simulation tools; and integrating these resources together using cloud microservices and an HPC backend connected via a REST API and a public WebUI. The community can easily extend BioSimulations to additional modeling approaches, modeling languages, simulation algorithms, and simulation tools. The core effort is currently focused on developing a framework to enable multi-component, multi-algorithmic simulations.
Presenting Author: Sotirios Panagiotou, Erasmus Medical Center

Authors: Panagiotou,, Sotirios (Erasmus Medical Center); (Fernàndez Santoro, Elías Mateo (Erasmus Medical Center); Strydis, Christos (Erasmus Medical Center)

Extending the reach of NeuroML toward full model capture

Slides:
https://drive.google.com/uc?export&id=11R5nSIyP9UIlTtODnvNnePrf5dnY4rqo

Abstract:
NeuroML is a very extensive specification for neuron and network models, that captures most model features in common use. However, it often happens that new models introduce new dynamics, new modes of interaction and experiment-specific rigging; as a result, in their full extent these models escape the NeuroMLv2 specification. Hence the desired dynamics must be realised through custom programming that is specific to, and intertwined with, a specific simulation program; this hampers interoperability and re-use.
In this work, we propose a versatile set of generalised model features, that can capture many commonly encountered - yet difficult to standardise - aspects of practical experiments. These features are described in a simulator-agnostic format that meshes with the NeuroML+LEMS ontology.
To demonstrate, we present a complete port of a complex BRIAN model used in active research, to NeuroML plus these new features. Furthermore, we provide proofs-of-concept for how many common types of simulated experiments are also covered by the new framework.

Presenting Author: Jardine Bartholomew, University of Washington, USA

Authors: Jardine Bartholomew Sauro, Herbert (University of Washington); Smith, Lucian (University of Washington)

MakeSBML: A tool for converting between Antimony and SBML

Abstract:
For the SBML community we have an updated implementation: MakeSBML (https://sys-bio.github.io/makesbml), a client-side web application that makes translating between Antimony and SBML easy and comes with built-in BioModels

(https://www.ebi.ac.uk/biomodels) search capabilities to simply download and convert curated BioModels to Antimony.

Presenting Author: Adel Heydarabadipour

Authors: Adel Heydarabadipour, Herbert M. Sauro
Department of Bioengineering, University of Washington, Seattle, WA 98195, USA.

SBMLNetwork: A python package to visualize and edit SBML model networks

Abstract:
SBMLNetwork is a python package designed to visualize and edit SBML model networks. It employs an embedded force-directed autolayout algorithm to create aesthetically pleasing networks and utilizes SBML Layout and Render packages to store and retrieve visualization information of models. With its drawing tool and its comprehensive command-line API, users can generate and modify graphical representations of their SBML models, save the visual information of their model to SBML files or strings, and export rendered figures of their model in PDF, PNG, and JPG formats. The package is installable via pip ('pip install sbmlnetwork') and the source code is available on https://github.com/adelhpour/SBMLNetwork.

Posters

Presenting Author: Manuel Lera-Ramirez, University College London

Authors: Manuel Lera-Ramirez Bähler, Jürg (University College London)

ShareYourCloning: towards a FAIR standard for sequence provenance

Abstract:
Recombinant DNA technology is used in diverse research fields to generate new DNA sequences by combining fragments from existing DNA molecules. These engineered molecules are introduced into cells as self-replicating plasmids and can be used to modify their genomes.
Currently, no stablished standard exists to describe how DNA sequences are generated from its ancestors. Therefore, researchers document the provenance of plasmids, cell lines and strains using spreadsheets or text-based systems, which are necessarily inconsistent and differ between collections. Some proprietary tools keep track of the provenance of sequences, but they don't allow users to export this information in an open format.
ShareYourCloning aims to develop a data model to describe sequence provenance while developing web tools that leverage the model and allow researchers to plan and document their experiments. In this way, the data is captured through an intuitive interface in a

standardised interoperable format at the beginning of the data lifecycle. For now the data is exported as json, but in the near future it will be possible to export a sequence and its provenance in SBOL format, following SBOL's PROV-O guidelines.
You can try the web application at https://shareyourcloning.netlify.app

Presenting Author: V Lokesh, Indian Institute of Technology Bombay

Authors: V Lokesh Wangikar, Pramod

Metabolic modeling of novel fast growing cyanobacterium Synechococcus elongatus PCC 11801 for value added products

Abstract:
Greenhouse gases from chemical manufacturing industries contribute significantly to climate change and global warming, and their sequestration is imperative to reducing the carbon footprint. Cyanobacteria are photoautotrophic microorganisms capable of fixing CO2 from the environment via photosynthesis, and can be engineered to convert the CO2 to value-added products. Despite progress in synthetic biology and strain engineering technologies, the commercial applicability of cyanobacteria has been limited due to the slow growth rate. However, the freshwater strain Synechococcus elongatus PCC 11801 recently isolated from our lab is one of the fastest growing strain with tolerance to different environmental stresses, making it a potential cell factory. To streamline and guide strain engineering efforts for PCC 11801, we generated a genome scale metabolic model for the strain that can be utilized with flux analysis to generate in silico predictions of hotspots for genetic perturbation that would maximize a specific value-added product in its metabolic network. The model was reconstructed using the standard protocol of draft reconstruction generation, manual curation, model debugging and validation. The genome scale model

('iJB785') of a phylogenetically close model strain Synechococcus elongatus PCC 7942 was used as the starting point. The draft reconstruction of PCC 11801 was generated by first mapping the homologous genes with iJB785 through bidirectional best hit analysis of the genomes. The model reactions were manually curated to reduce redundancy, and verified with the BiGG, KEGG and BRENDA database. 287 novel reactions identified pertinent to PCC 11801 were added de novo into the model with the corresponding gene-protein-reaction association. Model debugging was performed through flux balance analysis, flux variability analysis, single gene deletion analysis, and single reaction deletion analysis. Overall, the metabolic model consisted of 1052 genes and 1130 reactions. The model was validated by gene essentiality analysis with a sensitivity and specificity score of 0.9223 and 0.5425 respectively. The metabolic robustness of the model was estimated to be 39% from the model, which validates the amenability of the strain for genetic engineering. The model predicted that the strain would convert CO2 to succinic acid at a higher production rate compared to other value-added products through the in silico predictions.

Presenting Author: Francis Chemorion, University of Pompeu Fabra

Authors: Francis Chemorion

Enhancing Computational Biology Reproducibility and Accessibility through a Docker-Based Registry for SBML and SED-ML

Abstract:
The reproducibility of computational models in biology is hindered by the variability in execution environments, leading to challenges in model sharing and reuse. This work introduces a novel Docker-based registry designed to distribute Systems Biology Markup Language (SBML) models and Simulation Experiment Description Markup Language (SED-ML) files, along with their auto-generated interfaces and solvers. By encapsulating these elements in Docker containers, we ensure consistent execution across different computing environments, thus addressing a significant barrier to reproducibility and accessibility. Our approach not only facilitates the seamless sharing of computational models but also promotes the adoption of standardized practices within the scientific community. An evaluation of the registry underscores its potential to improve usability, foster collaboration, and enhance the integrity of computational research in biology. This initiative aligns with efforts to standardize computational models and their dissemination, promising to advance the field of computational biology by making robust, reproducible research more attainable.

Presenting Author: Michael Blinov, UConn Health

Authors: Michael Blinov Moraru, Ion; Schaff, James; Loew, Leslie

Virtual Cell (VCell) Modeling and Simulation Framework

Abstract:
Virtual Cell (VCell, http://vcell.org) is an open-source platform (automatic installers for Windows, Mac OS and Linux) that provides powerful capabilities for kinetic modeling of cellular systems. It provides one-stop simulation shopping: deterministic (compartmental ODE or reaction-diffusion-advection PDE), stochastic reactions (several SSA solvers), spatial stochastic (reaction-diffusion with Smoldyn), hybrid deterministic/stochastic and network-free agent-based simulations.
Simulations can run on our remote servers from any low-cost laptop. Optionally, simulations can also be run locally without an Internet connection. Models and simulations can be accessed from anywhere, and models can be shared among collaborators or made publicly available.
If diffusion and spatial localization of molecular species can affect the biology, the geometric shapes of the membrane and volumetric compartments can be explicitly considered. Model geometries may be derived from idealized analytical expressions or from experimental 2D or 3D microscope images, or constructed using constructive solid geometry.

VCell supports rule-based modeling using both network generation (BioNetGen software) and network-free simulation (NFSim software). SBML, SED-ML, and COMBINE archive exchange are supported.

Presenting Author: Natasa Miskov-Zivanov, University of Pittsburgh

Authors: Natasa Miskov-Zivanov Emilee Holtzapple, Casey Hansen

MINUET: Automated and systematic verification to increase quality and long-term reuse of models

Abstract:
Although modeling is an important component of a research pipeline in biology, most often there is no systematic or standardized approach for quality assessment and annotation of models, reducing their trustworthiness and reuse potential. Moreover, most of the model design and documenting steps are still done manually. Creating useful and reliable models of cellular signaling requires thorough and careful information extraction, knowledge assembly, and comprehensive model verification and validation. Manual execution of these steps can be highly impractical due to their complexity and the detailed attention they require. The verification step, assessing whether the model structure is correct by finding support for all its elements and interactions, and the validation step, evaluation of model behavior against experimental observations and data, usually occur iteratively with model expansion before the model can be used to make predictions or explanations.
In this work, we introduce MINUET (Managing Interaction and Network (re-)Usability through Evaluation of Trustworthiness), a platform for automated network verification and curation.

MINUET is comprehensive, conducting model verification that relies on both knowledge and data, the vast published literature and publicly available data-bases; it is flexible as it can conduct both context-aware and context-independent verification; it is versatile, allowing users to conduct in-design verification during model creation, post-design verification of existing models, or a comparison of models to verify them against each other; finally, it is fast, due to its automated steps for retrieving and comparing interactions. MINUET contributes to `the four R's` in several ways: it evaluates and facilitates reliability and reusability of models and information; it strengthens the potential for reproducibility of predictions across models by identifying structural differences between them; and it assesses the replicability of outcomes and observations by collecting evidence from knowledge and data sources.

Presenting Author: Gaoxiang Zhou, University of Pittsburgh

Authors: Gaoxiang Zhou Sayed, Khaled (University of New Haven);
Luo, Haomiao (University of Pittsburgh);
Arazkhani, Niloofar (University of Pittsburgh);
Ahmed, Yasmine (University of Pittsburgh);
Holtzapple, Emilee (University of Pittsburgh);
Tang, Difei (University of Pittsburgh);
Bocan, Kara (University of Pittsburgh);
Telmer, Cheryl (Carnegie Mellon University);
Miskov-Zivanov, Natasa (University of Pittsburgh);

A flexible modeling and simulation framework for hybrid elements

Abstract:
To retain the benefits of both logical and discrete modeling, and to adapt to the varying precision or amount of the available information, we introduce here the DiSH2.0 (Discrete, Stochastic, Heterogeneous) modeling and simulation framework. In DiSH2.0, both logical and algebraic update functions can exist within a hybrid model, enabling optimal use of the retrieved element and interaction information, including qualitative and causal information, as well as quantitative information when available. Instead of encoding discrete element values

with Boolean variables, using discrete variables in DiSH2.0 allows elements within a single model to have any necessary number of discrete levels and simplifies model creation, verification, and use. Additionally, the models have `memory`, as element update rules include the previous value of an element when calculating its change. The use of regulation weights and other notations in update rules offers a way to account for varying effects across different regulators as well as different levels of a single regulator.
For these new hybrid models, we introduced spontaneous increase and decrease behaviors, different responses to the balance between negative and positive regulators, and varying increment/decrement amounts. These features offer the flexibility to increase model accuracy when precise information is available or to retain a more abstract level in the absence of such information. We implemented a simulator that supports these new features and is also compatible with our previous work and other discrete modeling tools. The simulator expands to hybrid models the suite of deterministic and stochastic (also known as synchronous and random asynchronous, respectively) simulation schemes that have been mainly used for logical models in the past.
We demonstrated on a T cell differentiation use case the accuracy and precision improvements provided by hybrid modeling and new modeling and simulation features. Using the BioRECIPE representation format, which is both human and machine readable and editable, DiSH2.0 is compatible with other tools enabling full automation of the modeling pipeline: interaction extraction from literature and databases, interaction filtering, assembly of new and extension of existing interactions, and model verification and validation.
The repository of DiSH 2.0 is hosted at https://github.com/pitt-miskov-zivanov-lab/DiSH and its documentation at https://melody-dish.readthedocs.io

Presenting Author: Maren Philipps, University of Bonn

Authors: Maren Philipps Schälte, Y., Fröhlich, F., Jost, P. J., Vanhoefer, J., Pathirana, D., Stapor, P., Lakrisenko, P., Wang, D., Raimúndez, E., Merkt, S., Schmiester, L., Städter, P., Grein, S., Dudkin, E., Doresic, D., Weindl, D., & Hasenauer, J.

pyPESTO: PYthon Parameter EStimation TOolbox

Abstract:
pyPESTO is a widely applicable and highly customizable toolbox for parameter estimation. It can be used as a parameter estimation pipeline for systems biology problems specified in SBML and PEtab. The interface to AMICI enables efficient simulation and sensitivity analysis of ordinary differential equation (ODE) models. Toolbox features include multi-start local optimization, profile computation and visualization. Notably, it accommodates parameter estimation with ordinal data, censored data and nonlinear-monotone data.

Presenting Author: Sotirios Panagiotou, Erasmus Medical Center

Authors: Sotirios Panagiotou Fernàndez Santoro, Elías Mateo (Erasmus Medical Center); Strydis, Christos (Erasmus Medical Center)

Extending the reach of NeuroML toward full model capture

Abstract:
NeuroML is a very extensive specification for neuron and network models, that captures most model features in common use. However, it often happens that new models introduce new dynamics, new modes of interaction and experiment-specific rigging; as a result, in their full extent these models escape the NeuroMLv2 specification. Hence the desired dynamics must be realised through custom programming that is specific to, and intertwined with, a specific simulation program; this hampers interoperability and re-use.
In this work, we propose a versatile set of generalised model features, that can capture many commonly encountered - yet difficult to standardise - aspects of practical experiments. These features are described in a simulator-agnostic format that meshes with the NeuroML+LEMS ontology.
To demonstrate, we present a complete port of a complex BRIAN model used in active research, to NeuroML plus these new features. Furthermore, we provide proofs-of-concept for how many common types of simulated experiments are also covered by the new framework.

Presenting Author: Dilan Pathirana, University of Bonn

Authors: Dilan Pathirana

PEtab v2 Roadmap

Abstract:
PEtab is a standardized file format for specifying parameter estimation problems for dynamic models based on Ordinary Differential Equations [1]. The interoperable format is currently supported by 11 different tools [2], enabling users to benefit from standardized parameter estimation across frameworks based in Python, Julia, R, MATLAB, C++, or GUIs. This poster will complement the breakout session, by outlining the current developments in the format, with a view towards the next version of PEtab. This includes unit specification, reproducible result storage, model selection, and time-course specification.
[1] ""PEtab—Interoperable specification of parameter estimation problems in systems biology"" https://doi.org/10.1371/journal.pcbi.1008646
[2] https://github.com/PEtab-dev/petab#petab-support-in-systems-biology-tools