Integer Linear Programming in Computational and Systems Biology: Tutorial

Dan Gusfield

gusfield@cs.ucdavis.edu

Department of Computer Science, University of California, Davis
Davis, California

CCS CONCEPTS

• **Applied computing** → **Computational genomics; Computational genomics; Systems biology; Bioinformatics; Decision analysis.**

KEYWORDS

integer linear programming, computational biology, systems biology

ACM Reference Format:

Dan Gusfield. 2019. Integer Linear Programming in Computational and Systems Biology: Tutorial. In *10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB '19), September 7–10, 2019, Niagara Falls, NY, USA*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3307339.3343179>

1 GOAL OF THE TUTORIAL

The goal of the tutorial is to introduce and detail *modeling* and *solving* of real problems in computational and systems biology using integer linear programming. This is in contrast to goal of explaining integer programming theory, or how integer programming solvers work.

2 WHY INTEGER PROGRAMMING IN BIOLOGY

Integer Linear Programming is a versatile modeling and optimization technique that was developed for complex planning and operational decision making. However, it is increasingly used in computational biology in non-traditional ways, most importantly and inventively as a computational tool and language to model and study biological phenomena, to analyze biological data, and to extract biological insight from the models and the data.

Integer linear programming is often very effective in solving *instances* of biological problems on realistic data of current importance, even for hard computational problems that lack a worst-case efficient solution method. For any given problem instance, highly engineered ILP solvers can be

used to solve the ILP formulation. The improvement of the best ILP solvers has been spectacular, with an estimate that (combined with faster computers) benchmark ILP problems can now be solved 200-billion times faster than twenty-five years ago. The effectiveness of the best modern ILP solvers on problem instances of importance in biology opens huge opportunities – broad exploitation of integer programming could have a truly transformative effect on computation in biology and perhaps medicine.

3 TUTORIAL OVERVIEW

The tutorial will start with an introduction for people with no prior background in linear programming or integer linear programming. We will discuss several problems, models and ILP solutions selected from network biology, phylogenetics, and RNA folding. The ILP formulations will illustrate several different ILP “idioms” and their use. Finally, after studying several ILP examples in computational biology, we will discuss what is distinctive about the use of integer programming in computational biology, compared to more traditional uses of integer programming.

We will illustrate some concepts using a commercial ILP solver from Gurobi Optimization, to solve specific ILPs that we formulate.

Material for this tutorial is partially extracted from [1] and [2].

ACKNOWLEDGMENTS

This work was partially supported by NSF grant 1528234.

4 ONLINE RESOURCES

Python and Perl programs related to this tutorial can be downloaded from:

www.cambridge.org/9781108421768

REFERENCES

- [1] Dan Gusfield. 2019. *Integer Linear Programming in Computational and Systems Biology: An entry-level text*. Cambridge University Press.
- [2] Dan Gusfield. 2019. Integer Linear Programming in Computational Biology: Overview of ILP and New Results for Traveling Salesman Problems in Biology. In *Bioinformatics and Phylogenetics, Computational Biology 29*, Tandy Warnow (Ed.). Springer Nature.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ACM-BCB '19, September 7–10, 2019, Niagara Falls, NY, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6666-3/19/09.

<https://doi.org/10.1145/3307339.3343179>