Introduction

As the amount of biological data grows, the task of understanding existing data becomes increasingly important, and this is largely a task best undertaken by computer science. This book is for the many curious souls who are coming into biology from backgrounds in computer science, especially the fields of information retrieval, natural language processing, and/or machine learning.

One major difference between biology and computer science is that in computer science, the world we explore is in large part our own creation, and a large part of what we do is make our creation understandable by finding useful abstractions, and then building more complex things by combining these abstractions together. For example, a deterministic finite state machine is a useful abstraction for computations that process discrete inputs sequentially with limited memory—we study this, and study stack data structures, and then study the result of combining them to make a push-down automaton. These abstractions might be compromised when we optimize our systems for performance, but they are rarely abandoned completely, because comprehensibility, elegance, and simplicity are practically important for systems that must be maintained and improved by humans.

In contrast, biology doesn't lend itself to clean and comprehensible abstract models: evolution relentlessly marches toward improved performance without worrying much about simplicity. Even the "simplest" forms of life are seemingly endless in their unique complexities, and almost every general statement about how organisms function comes with an asterisk. And unlike in computer science, the details that underlie the complexity of the real

W. W. Cohen, C. K. Cohen, *A Computer Scientist's Guide to Cell Biology*,
https://doi.org/10.1007/978-3-031-55907-5_1

systems are not something we can or should ignore, hoping they will be cleaned up in the next version—instead, the awkward details are, collectively, the real subject of the science of biology.

For the purposes of this book, we have broken the field down into three parts:

- **Biological mechanics** are the actual nitty-gritty details of how things work at the cellular level—protein pathways, chloroplasts, and so on. This is the typical focus of introductory biology classes and textbooks, and you would correctly suppose this is the essence of what biologists actually study. However, it's a surprisingly small part of what biologists write and talk about.
- **Experimental methods**, on the other hand, *are* what biologists spend most of their time talking about. If you pick up a typical biology paper, the actual *conclusions*, the newly discovered details about how these systems function, are compact enough to be laid out in the abstract.

As you read through this book, you'll find that it's mostly about methods. Biologists spend most of their word count talking about how they conducted their experiments, how cells were cultured, and what assays were run and a host of other details. The results, in isolation, tell you very little—the only way to tell the difference between good research and bad research is to examine how data was initially gathered. But to an outsider, that can be far from a simple task. The language of biology is rich, detailed, and almost impenetrable to the average layperson; learning its intricacies is as important as learning about biological mechanisms or experimental techniques.

- **Language and nomenclature** can be considered a "part" of biology in its own right. Without spending at least a little bit of time learning how to speak it, this book would be pretty useless.

If you like, you can think of biology as a journey to some strange, exotic land. The inhabitants speak a strange and often incomprehensible language, the customs and practices may be like nothing seen before, and even the most basic of tasks appear completely alien. With that in mind, our goal is to provide a short introduction to the three core aspects of cell biology—a travel guide, to continue the previous metaphor, focusing on high-level principles, and relating as much as possible to familiar concepts from computer science.

Consequently, in this book, we will gloss over some concepts and oversimply others, setting aside many otherwise-fascinating theories and details. Biology is fractal; no matter how deep you look, there is always another layer of complexity. For a more comprehensive background on biology, there are many excellent textbooks, written by people far more qualified—the last chapter of this book will introduce several of our favorites.