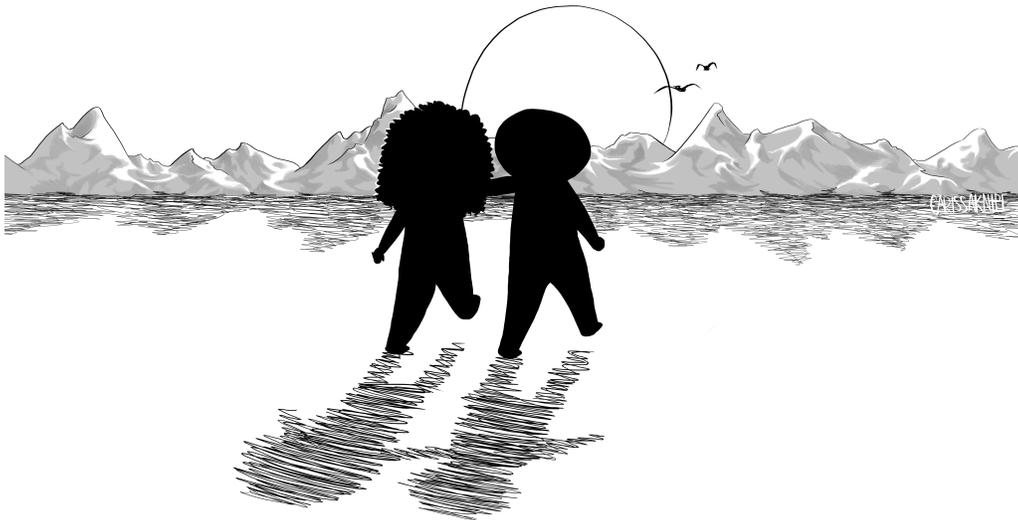


# 12 Looking Forward

---



*In which our heroes reflect on where they have been, and look forward to where they will travel in the future.*

12-1

## 12-2 Looking Forward

We have always held to the hope, the belief, the conviction, that there is a better life, a better world, beyond the horizon.

---

Franklin D. Roosevelt (1882–1945)

radio address on hemisphere defense from Dayton, Ohio (1940)

This book has introduced the mathematical foundations of computer science—the conceptual building blocks of, among other things, the large, complex computational systems that have become central aspects of our daily lives, some of which have already genuinely and meaningfully improved the world in their own unique ways, profound and small. Understanding and reasoning about these fundamental building blocks is necessary for you to understand, develop, and evaluate the key ideas of these many new applications of computer science, and introducing these foundations has been the underlying goal of this book.

We’ve seen introductions to a wide variety of those applications as we’ve gone along, but perhaps the single best example is the idea of paired kidney transplants (p. 11-71). The technical background that enables the paired-transplant system uses foundational concepts from at least five chapters of this book and, far more importantly, saves literally hundreds to thousands of lives per year. There are plenty of other examples: improved forecasting and communication systems give earlier warnings about an impending natural disaster; tools facilitate rerouting unused restaurant food to food banks instead of landfills; screen readers allow visually impaired people to access information that would otherwise have been unavailable except at great expense.

There are two requirements for these interventions to have worked as well as they have: first, having the underlying idea for the system, and, second, implementing it correctly. It’s possible, of course, to conceive of a computational approach to a problem that, if solved properly, could have this same kind of profoundly positive impact on the world—but to execute that solution buggily. The Therac-25 system (p. 4-82) was a paradigmatic example of this phenomenon: a medical device built to treat tumors that, because of a particular fault that occurred when the machine’s operator was too fast a typist, ended up killing patients instead of helping them to heal.

But if it’s bad to implement good things buggily, it’s worse to do bad things well. There is simply no circumstance in which ransomware or a tool for phishing that attempts to steal someone’s financial credentials, for example, should be built. Aside from these very rare cases in which someone conceives of a computational idea that’s unmitigatedly awful, though, nearly every idea has both benefits and costs, and the ethically complex questions are all about the tradeoffs between them.

Bitcoin—a digital currency built around cryptography using what’s called a “blockchain”—allows individuals to form a fully decentralized financial system, without needing to trust (or fund) big banks. And yet the bitcoin “mining” and transaction-verification processes by which the system works are extremely energy intensive, and environmental consequences are potentially large.

More generally, cryptographic systems like Bitcoin, Tor (p. 5-22), and RSA encryption (Section 7.5) are intended to allow communication and transactions that are fully private. And yet that privacy guarantee also implies untraceability even if the content is about the trafficking of money or weapons or people.

Meanwhile, systems based on image processing can provide security on a smart phone, or take readings of blood oxygenation levels for rapid diagnosis, or personalize information of particular interest to *you* as you move through the world. And yet those same systems can make stalking or government surveillance easier, and the systems that have actually been deployed have tended to have gender- or race-biased outcomes. (For example, automated systems that seek to caption images have been shown to exacerbate gender stereotyping even beyond the disparities present in training data [133]; Black men have been arrested for crimes that they did not commit, based on erroneous facial recognition [54]; and the accuracy of commercially released image-processing systems is highest for lighter-skinned male faces and lowest for darker-skinned female faces [24].)

It would take a whole other book to do justice to the question of how to appropriately weigh the benefits and costs of potential computational innovations. But there are a few general categories of questions that are helpful to consider as you're thinking about these tradeoffs. Who would be served well by the system that you're thinking about building? Who would be served less well? Who would it ignore? Who might it harm? What are the risks: what could go wrong, and what are the potential unintended consequences of deploying the system? (For another way of thinking about these concerns: the Association for Computing Machinery, one of the main professional societies for computer scientists, has published a Code of Ethics and Professional Conduct, <https://www.acm.org/code-of-ethics>, that articulates the underlying principles as statements of responsibility, rather than as questions to consider.)

---

In the mid-1960s, the German-American computer scientist Joseph Weizenbaum (1923–2008) developed a computer system that could carry out a psychiatrist-like conversation with a human user [129]—a

```
ELIZA: I am the psychotherapist. Please, describe your problems.
ME: I am worried about the future of the world.
ELIZA: Are you worried often?
ME: Yes.
ELIZA: I see... Well, why do you say this is so?
ME: I think that technology can be used for a great deal of good and it often is, but sometimes it's
used for malice and sometimes it's developed to try to do something good but is implemented in a way
that makes the world worse.
ELIZA: Is it because sometimes it's used for malice and sometimes it's developed to try to do
something good but is implemented in a way that makes the world worse that you came to me?
[...]
```

**Figure 12.1** A conversation with a version of ELIZA (the version embedded in a recent version of the program Emacs).

## 12-4 Looking Forward

1960s version of a chatbot. The system, called ELIZA, was rather technically simple: it “knew” a collection of scripted patterns of communication and some grammatical rules for converting a user’s statement into a question—as in responding to *I wish I were more social* with *Do you really think you wish you were more social?*. (Technically, it used the kinds of grammar-based tools discussed on p. 5-56.) See Figure 12.1 for a sample dialogue with ELIZA.

After Weizenbaum built ELIZA, he observed people using it—and, surprisingly often, confiding deeply personal information in conversations with it—and he found himself troubled by the ways in which humans acted as though it were truly intelligent. This experience started Weizenbaum onto a more philosophical track in thinking about artificial intelligence specifically, and computing in general, and eventually turned him into a skeptic of the expanding role of computers in life. In 1976, he published a book entitled *Computer Power and Human Reason* [130], expressing some of this skepticism about what role computers play in society, in ways that feel prescient today:

[Another] kind of computer application that ought to be avoided, or at least not undertaken without very careful forethought, is that which can easily be seen to have irreversible and not entirely foreseeable side effects. If, in addition, such an application cannot be shown to meet a pressing human need that cannot readily be met in any other way, then it ought not to be pursued. (p. 270)

(In the latter category, Weizenbaum—who, with his German Jewish family, was forced to leave Nazi Germany in 1936—talks about speech recognition as an example, pointing out the ways that government surveillance would become so much easier in the presence of automatic speech recognition. He’d have felt even more strongly about facial recognition, presumably.)

Earlier in the book, Weizenbaum describes what he believes to be appropriate boundaries for the role of computation in our lives:

What I conclude here is that the relevant issues are neither technological nor even mathematical; they are ethical. They cannot be settled by asking questions beginning with “can.” The limits of the applicability of computers are ultimately statable only in terms of oughts. What emerges as the most elementary insight is that, since we do not now have any ways of making computers wise, we ought not now to give computers tasks that demand wisdom. (p. 227)

---

If this book has done its job, you are by now equipped with the key mathematical foundations necessary to create new ideas, new insights, new applications of computer science. There are so very many ways in which the world can be made better through those new contributions. (And some that make it worse: as Weizenbaum wrote, “a person falling into a manhole is rarely helped by making it possible for him to fall faster or more efficiently.”) May you, and may we all, choose to build the things that make the world better.