

# THE PUZZLE HAS A POINT

Charles Kolozsvary

*To Nick:  
For bringing to my attention the threat of certain crustaceans*

# Preface

Each month, a quantitative trading firm called Jane Street shares a puzzle so abstract and complicated that you're left wondering who on earth could have come up with it.<sup>1</sup> The answer: some nerd genius making way more money than you might think possible as someone whose job it is to solve puzzles like this every day. Putting aside whether you think a dense math problem like the one in the last FN really deserves to be called a puzzle,<sup>2</sup> its worth investigating how one might solve them. there's something worth investigating here, whether it's how a proprietary trading firm like Jane Street is so successful or how you'd go about solving these tricky problems.

Maybe to the disappointment of a few, what follows will delve into the latter question, presenting solutions to a handful of Jane Street's puzzles.<sup>3</sup> But we shouldn't despair: math has a lot going for it, at least in this author's opinion. It might be the most powerful, dynamic, and fascinating subject of all time.<sup>4</sup> However, like anything that's worth doing, math requires effort, for the experienced and uninitiated alike. The former<sup>5</sup> may wish to skip directly to Chapter 2, and the latter needn't worry: The presentations are meant to be understood by anyone older than, say, fifteen.<sup>6</sup>

By the end of this unorthodox tour of mathematics, we'll have gained a handful of tools that will help us reach even greater heights and depths of the subject. Or, at the very least, we'll have solved some interesting puzzles. And, if you're still reading, you probably agree, as Jane Street also says, that "solving puzzles just feels great."

---

<sup>1</sup>Example: Consider 3-space partitioned into a grid of unit cubes. Given a fixed positive real number  $D$  and a random line segment of that length, what value  $D$  maximizes the probability that the two endpoints of the segment lie in orthogonally adjacent unit cubes, and what is this maximal probability?

<sup>2</sup>(which it does)

<sup>3</sup>For the time being, only a presentation of the August 2024 puzzle, *Tree-edge Triage*, is included.

<sup>4</sup>To demonstrate this claim would be to teach all of mathematics, and no one has the time or skill to do this. But as the saying goes, "don't let perfect be the enemy of good." We need to start somewhere, so we start here, now, with some puzzles.

<sup>5</sup>(those already versed in basic set and probability theory)

<sup>6</sup>This is not meant to make anyone feel discouraged, of course, if some ideas do not come easily; the intention of such a statement is rather that anyone (yes, anyone) can eventually understand what we're doing.

*Ask, and it shall be given you; seek, and ye shall find; knock, and it shall be opened unto you — Matthew 7:7*

*If you can't explain it to a six year old, you don't understand it yourself — Albert Einstein*

# Chapter 1

## Preliminaries

The “royal we” in the preface is in fact accurate: this section includes exercises, and you are actually expected to do them. But don’t worry;<sup>1</sup> the exercises are short and simple and their solutions are in Appendix A.<sup>2</sup>

### 1.1 The Beginning

*Natural numbers were created by God, everything else is the work of men* — Kronecker (1823–1891)

You might not think about it like this, but the very first thing you learned in math was how to count — e.g., there are three apples over here and two over there, and if we put them together, there are five apples.<sup>3</sup>

Existing in thought or as an idea but not having a physical or concrete existence

The numbers involved in this basic form of counting are aptly<sup>4</sup> called the *natural numbers*. One is a natural number, and so is two and three and four and so on.<sup>5</sup>

---

<sup>1</sup>Be happy.

<sup>2</sup>A final word of advice: give each exercise an honest attempt before consulting its solution; you’ll learn more this way, and you’ll learn faster.

<sup>3</sup>INTERPOLATION: THE THING, NOT IN ITSELF

People lead disparate lives, and there is an incomprehensible range of possible experience. So it should be no surprise that there are many schools of thought when it comes to putting (or, rather, attempting to put) slippery philosophical musings precisely. The we investigate here is *abstraction*. The Oxford English Dictionary (OED) says that there exists a word called ‘abstraction’ which plays a crucial role in mathematics and It’s worth appreciating, too, the following leap in meaning that is often taken for granted: there is something *the same* about a pair of apples, rhinos, rocks, zooxanthellae, etc. That *thing* is, in itself, ‘the number two’, and to give it the respect and distinction it deserves, we refer to it with a symbol that should remind us that we’re interacting with something unusual: ‘2’ (the unusual thing being *abstraction*).

<sup>4</sup>(monists would disagree)

<sup>5</sup>9001 is a natural number, too.

But, as almost everyone already knows, these aren't the only numbers. Some numbers are . . . *unnatural*, if history is any measure. Zero and negative numbers, for instance, were adopted well after the naturals. And it isn't too hard to imagine why: neither of them can be thought of in the same straight-forward, tangible way as a natural number — good luck pointing at zero or negative three apples. We need to think about these numbers a little differently, i.e., we need to recognize that each is a unique leap in abstraction: zero represents the absence of a quantity rather than its precise multiplicity (as with the naturals); and a negative number represents a *reduction* in number rather than a number outright.

You might think a noun like “abstraction” overcomplicates a simple idea like zero or a negative number. You'd be wrong, however, or there would only be at most a disagreement on definitions, not real, capital ‘M’ Meaning.

And there are even more numbers! We notice that some things can be considered as parts of a whole rather than whole things themselves: you can have half of an apple, or a quarter, or three seventh's and so on. But the trickery here is assuming that any kind of whole exists that can be considered in parts as opposed to everything being

## 1.2 Sets

A set is a collection of distinct items or “elements.” Sets are written by listing their elements inside curly braces and can be represented by variables. Here are some example sets:

$$(1) \{terrier, poodle, Godzilla\},$$

$$(2) A = \{a, \dots, z\},$$

$$(3) \mathbb{N} = \{1, 2, 3, \dots\}.$$

Ellipses ‘...’ indicate implied elements that are tedious or impossible to write explicitly, as in items (2) and (3), respectively. These examples also show that sets can be finite or infinite.<sup>6</sup> The number of elements in a set  $S$ , known as its cardinality, is denoted  $|S|$ . For example,  $|A| = 26$ .

**Exercise 1.1.** What is  $|\{3, 3, 1, 2, 1, 3\}|$ ?

### 1.2.1 The Cartesian product

The Cartesian product of two sets  $S$  and  $T$  is the set of all ordered pairs with the first element from  $S$  and the second from  $T$ , denoted  $S \times T$ . For example,

$$\{1, 2\} \times \{a, b, c\} = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}.$$

---

<sup>6</sup>In fact, there are—in an abstract but rigorous sense—infinately many and ever-growing kinds of infinities. But this is beyond our scope. For more on this, look up Cantor's Theorem.

**Exercise 1.2.** For finite  $S$  and  $T$ , what is  $|S \times T|$  in terms of  $|S|$  and  $|T|$ ?

The title of this subsection is a set that comes up a lot in mathematics. Putting aside rigor, the real numbers are those that can be written as an infinite string of digits, repeating or otherwise, such as  $3.1415926\cdots = \pi$  or  $-0.0000000100003$ . These numbers

This brief description misses some important and fascinating details that the curious reader might want to look into by searching “transcendental vs. algebraic numbers.”

### 1.2.2 Bounds of sets

A set may have lower (or upper) bounds when its elements can be ordered. For example, consider the set  $H = \{0, 1, 2, 3\}$ . The number  $-100$  is a lower bound of  $H$  because it is less than or equal to every element of  $H$ . However, there are greater lower bounds, such as  $-99$ ,  $-98$ , and so on.

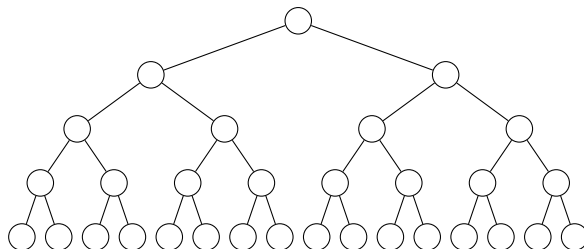
**Exercise 1.3.** What is the greatest lower bound of  $H$ ? Why?

We call the greatest lower bound of a set its *infimum*. The minimum of a set is similar, it is the element that is less than every other element.

**Exercise 1.4.** Are the infimum and minimum of a set always the same? If they aren't, give an example.

## 1.3 Binary Trees

Figure 1.1: the circles are called nodes, and the lines connecting them are called edges. The topmost node is called the root, and the bottom nodes are called leaves. The root has two descendants (hence the adjective “binary”), each of which has two descendants of their own and so on. The depth of a tree is the number of edges between the root and any of its leaves; this tree's depth is four.



## 1.4 Probability

An event is *random* when it cannot be predicted with complete certainty. The probability of an event  $E$  is its likelihood of occurring, denoted  $P(E)$ . For example, the probability of rolling a 6 on a standard six-sided die (denoted “d6”) is  $1/6$ .

If an event is guaranteed to happen, its probability is 1. So for any event  $E$ ,  $0 \leq P(E) \leq 1$ .

**Exercise 1.6.** Why? *Hint:* This is not a trick question.

For convenience, we can denote random events with variables. For example, let  $D_n$  be the roll of a  $dn$  (a standard  $n$ -sided die); then, again,  $P(D_6 = 6) = 1/6$ .

However, finding the probability of an event can be challenging when the context is unclear (e.g., what is the probability you think of the number 42?). Therefore, it’s helpful to keep in mind the set of all possible outcomes (relevant to an event), called the *sample space* and denoted  $\Omega$ . For example,  $\Omega = \{1, 2, 3, 4, 5, 6\}$  when rolling a d6: each number represents the outcome that the d6 rolls that value.

When  $\Omega$  is finite and each event is equally likely, finding probabilities usually reduces to counting outcomes. Specifically:

$$P(E) = \frac{\text{number of outcomes where } E \text{ occurs}}{\text{total number of outcomes}}. \quad (1.1)$$

**Exercise 1.7.** What is  $P(D_6 \text{ is prime})$ ?<sup>7</sup>

If outcomes are not equally likely, we can modify the sample space by assigning each event a number of outcomes proportional to its likelihood. For instance, if we have a weighted d4 where each number  $x$  comes up  $(10 \cdot x)\%$  of the time, this would be equivalent to a d10 with the number 1 on one face, 2 on two faces, and so on.

**Exercise 1.8.** Imagine two coins, each with sides labelled 1 and 2. One coin,  $F$ , is fair,<sup>8</sup> while the other one,  $M$ , is half-magic: if  $F = 1$ , then  $M = 1$ ; otherwise,  $M$  is fair. You flip  $F$ , then  $M$ . Describe a weighted die whose outcomes match the random variable  $F + M$ . *Hint:* apply equation (1.1) and the preceding paragraph.

**Independence.** Two random events are *independent* if neither gives information about the likelihood of the other. For example, the first roll of a d6 gives no information about its second roll, so the two rolls are independent, unlike  $M$  and  $F$  in Exercise 1.8. Independent events are almost always easier to reason about than non-independent (i.e., dependent) events.

<sup>7</sup>A prime number is only divisible by 1 and itself. But the first prime number is 2, not 1, because excluding 1 simplifies many theorems about primes.

<sup>8</sup>(meaning each outcome is equally likely)



**Exercise 1.9.** Why? *Hint:* compare the probability that the 100th roll of a d6 is 6 with the probability that the 100th roll of a *special* d6 (identical to a standard one, except if you ever roll more than 20 of a number, it is replaced by the number 1) is 6.

**Multiple events.** Events are usually composed of subevents. There are three fundamental operations that create new events from existing ones.

- (1) Negation: The event where  $E$  does not occur is written  $\neg E$ .
- (2) Intersection: The event where both  $E$  and  $F$  occur is written  $E \cap F$ .
- (3) Union: The event where  $E$  or  $F$  (or both) occur is written  $E \cup F$ .

**Exercise 1.10.** If  $P(E) = p$ , what is  $P(\neg E)$ ?

**Exercise 1.11.** Suppose events  $E$  and  $F$  are independent. Can you guess what  $P(E \cap F)$  is? *Hint:* Recall the Cartesian product.

**Exercise 1.12.** Let  $E$  and  $F$  be any two events. What is  $P(E \cup F)$  in terms of  $P(E)$ ,  $P(F)$ , and  $P(E \cap F)$ ? *Hint:* Consider the event of rolling an even or prime number with a d6, and draw a Venn diagram.

## 1.5 Polynomials

**In general.** A polynomial is an expression made up of indeterminants where the only operations involved are addition, subtraction, and multiplication. For example  $x^2 - x - 1$  is a polynomial made up of three terms and one indeterminant. This is a much clearer way of representing “the product of a value with itself, subtracted by itself and 1,” wouldn’t you say?<sup>9</sup>

The degree of a polynomial is the highest power of any indeterminant. For example,  $x^3 - 7x + y$  has degree 3. A polynomial is called quadratic if its degree is 2.

**Zeros and quadratic polynomials.** The zeros of a polynomial are, unsurprisingly, the values of the variable(s) that make the polynomial equal to zero. For example, the zeros of  $x^2 - 5x + 6$  are  $x = 3$  and  $x = 2$ . Finding the zeros of a linear (degree 1) polynomial is trivial, but less so for a quadratic.

**Exercise 1.13.** Find a formula for the zeros of a quadratic  $p(x) = ax^2 + bx + c$  in terms of its coefficients,  $a$ ,  $b$ , and  $c$ .<sup>10</sup> *Hint:* Can you turn  $p(x)$  into a form that looks like  $(x + d)^2 = e$ ?<sup>11</sup>

<sup>9</sup>We are lucky: mathematics was largely written in words before René Descartes popularized symbolic notation in his “La Géométrie,” published in 1637.

<sup>10</sup>The coefficients  $a$ ,  $b$ , and  $c$  are not to be confused with indeterminants: they represent only one possible value, unlike indeterminants that represent several possible values. In the earlier polynomial,  $a = 1$ ,  $b = -5$  and  $c = 6$ .

<sup>11</sup> $d$  and  $e$  are not indeterminants; they are particular values, like  $a$ ,  $b$ , and  $c$ .



## Chapter 2

# Tree-edge Triage

Aaron and Beren are playing a game that takes place within an infinite binary tree. At the start of the game every edge in the tree is independently labeled  $A$  with probability  $p$  and  $B$  otherwise.<sup>1</sup> Both players are able to inspect all (infinitely many) of these labels ... somehow. Then, starting with Aaron at the root, the players alternate turns moving a shared token down the tree (each turn the active player selects from the two descendants of the current node and moves the token along the edge to that node). If the token ever traverses an edge labeled  $B$ , Beren wins the game. Otherwise, Aaron wins.<sup>2</sup>

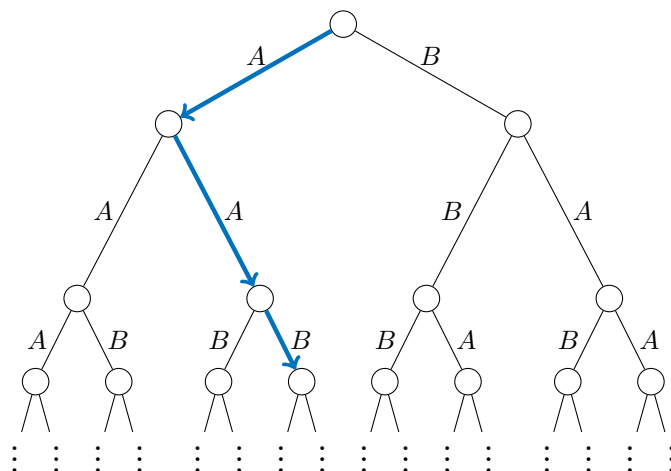


Figure 2.1: An example game: Aaron chooses to go left to avoid immediate defeat. However, after Beren goes right, Aaron is forced to choose one of two  $B$ -edges, and Beren wins.

<sup>1</sup>On average, then,  $p$  represents the proportion of edges in the tree that are labeled  $A$ .

<sup>2</sup>This is the August 2024 puzzle, FYI, that can be found at <https://www.janestreet.com/puzzles/tree-edge-triage-index/>

**Q:** What is the infimum of the set of probabilities  $p$  for which Aaron has a nonzero probability of winning the game?

This question is kind of dense for the sake of rigor that isn't that important right now. So here's another way to put it: what is the smallest average proportion of edges in the tree labelled  $A$  such that Aaron still has a chance of winning?

## Chapter 3

# Solution

Rather than tackling the entire infinite problem right away (which could be overwhelming), we first consider finite games restricted to finite binary trees. For these games we define the following events:

- Let  $X_n$  denote Aaron wins a game that starts on **his** turn and lasts  $n$  rounds.
- Let  $Y_n$  denote Aaron wins a game that starts on **Beren's** turn and lasts  $n$  rounds.
- Let  $A_e$  denote edge  $e$  is labelled  $A$ .

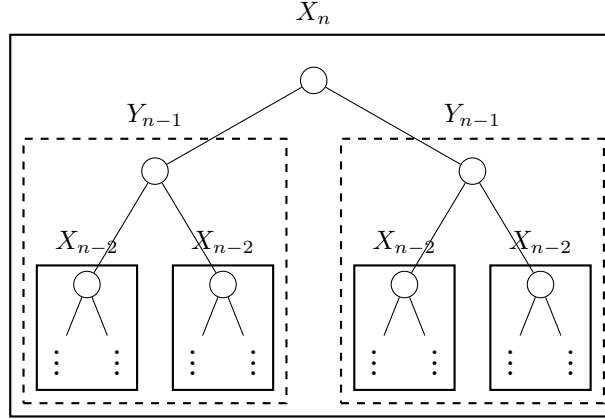
Now let's find some probabilities. The simplest one is  $P(A_e)$ , which, by definition, is  $p$  for any edge  $e$ . Next, it is clear that  $X_1$  occurs when at least one of the two edges from the root,  $e_1$  or  $e_2$ , is labeled  $A$ . Thus,

$$\begin{aligned} P(X_1) &= P(A_{e_1} \cup A_{e_2}) \\ &= P(A_{e_1}) + P(A_{e_2}) - P(A_{e_1} \cap A_{e_2}) \\ &= 2p - p^2. \end{aligned}$$

Similarly,  $Y_1$  occurs when both edges from the root are labelled  $A$  (otherwise, Beren wins). Thus,

$$P(Y_1) = P(A_{e_1} \cap A_{e_2}) = p^2.$$

Why stop with  $P(X_1)$  and  $P(Y_1)$ ? Let's try to find  $P(X_2)$ . We could apply the same reasoning that got us  $P(X_1)$ , noting that  $X_2$  occurs when either  $e_1$  and its subsequent edges are labelled  $A$  or  $e_2$  and its subsequent edges are labelled  $A$ . But this approach becomes unwieldy as we consider more and more rounds. It would be nice if we could express  $P(X_2)$  using what we already found—namely,  $P(X_1)$  and  $P(Y_1)$ . And indeed, we can! The key is to realize that our events are composed of identical sub-events (see Figure 3.1), thanks to all the edges being labelled independently.

Figure 3.1: The sub-events of  $X_n$ 

Thus,  $X_2$  occurs if  $e_1$  is labelled  $A$  and Aaron wins the rest of the game after traversing  $e_1$  ( $Y_1$ ) or  $e_2$  is labelled  $A$  and Aaron wins the rest of the game after traversing  $e_2$  ( $Y_1$ ). That is,

$$\begin{aligned}
 P(X_2) &= P((A_{e_1} \cap Y_1) \cup (A_{e_2} \cap Y_1)) \\
 &= P(A_{e_1} \cap Y_1) + P(A_{e_2} \cap Y_1) - P((A_{e_1} \cap Y_1) \cap (A_{e_2} \cap Y_1)) \\
 &= 2pP(Y_1) - p^2P(Y_1)^2.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 P(Y_2) &= P((A_{e_1} \cap X_1) \cap (A_{e_2} \cap X_1)) \\
 P(Y_2) &= p^2P(X_1)^2.
 \end{aligned}$$

But wait a second. Was there anything stopping us just now from writing  $X_n$  instead of  $X_2$  and  $Y_{n-1}$  instead of  $Y_1$  or  $Y_n$  instead of  $Y_2$  and  $X_{n-1}$  instead of  $X_1$ ? Absolutely not! This means that

$$P(X_n) = 2pP(Y_{n-1}) - p^2P(Y_{n-1})^2 \text{ and} \quad (3.1)$$

$$P(Y_n) = p^2P(X_{n-1})^2. \quad (3.2)$$

We now return to the infinite game. To answer question 2, we need to determine for which values of  $p$   $P(X_\infty) > 0$ . First, let's find  $P(X_\infty)$ . If we take the limit as  $n$  approaches infinity in (3.1) then substitute in the value of  $P(Y_n)$  from (3.2), we get

$$\begin{aligned}
 P(X_\infty) &= 2pP(Y_{\infty-1}) - p^2P(Y_{\infty-1})^2 \\
 &= 2p\left(p^2P(X_{\infty-2})^2\right) - p^2\left(p^2P(X_{\infty-2})^2\right)^2 \\
 &= 2p^3P(X_{\infty-2})^2 - p^6P(X_{\infty-2})^4.
 \end{aligned}$$

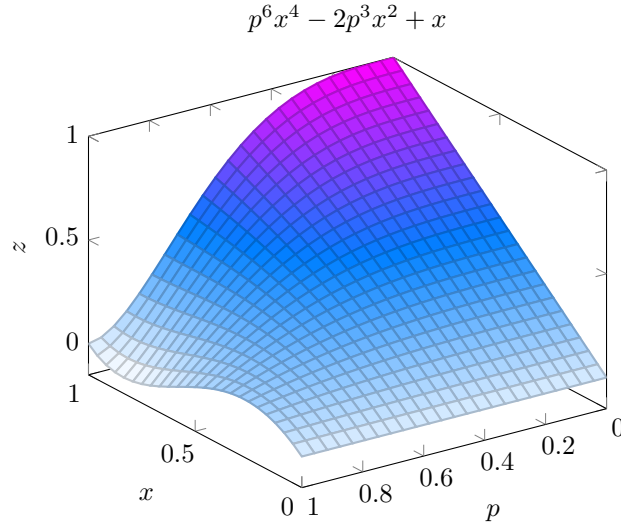


Figure 3.2: something

$P(X_\infty) = P(X_{\infty-2})$ , and we still don't know what  $P(X_\infty)$  is, so let it be  $x$ . Then we have  $x = 2p^3x^2 - p^6x^4$  or, after rearranging terms,

$$x(p^6x^3 - 2p^3x + 1) = 0.$$

We ignore the trivial solution  $x = 0$  and set our sights on  $p^6x^3 - 2p^3x + 1 = 0$ . Let's remind ourselves what this equation represents. The possible values of  $x$  that satisfy this equation are the possible values of  $P(X_\infty)$ . This is why we ignored  $x = 0$ ; we already know that Aaron's probability of winning can be 0 (set  $p = 0$ , for example).

The discriminant for a cubic polynomial of the form

$$x^3 + ax^2 + bx + c$$

is

$$a^2b^2 + 18abc - 4b^3 - 4a^3c - 27c^2.$$

Therefore

$$\begin{aligned} \Delta &= 0 + 0 - 4b^3 - 0 - 27c^2 \\ &= -4(-2p^{-3})^3 - 27(p^{-6})^2 \\ &= 2^5p^{-9} - 27p^{-12}. \end{aligned}$$

Setting  $\Delta = 0$  and solving for  $p$ , we find that

$$2^5 p^{-9} = 27 p^{-12}$$

$$2^5 p^3 = 3^3$$

$$p^3 = 3^3 \cdot 2^{-5}$$

$$p = 3 \cdot 2^{\frac{-5}{3}},$$

which is our final answer.



# Appendix A

## Answers to the Exercises

**Answer (1.10).** If  $P(E) = p$ , then  $P(\neg E) = 1 - p$ .

**Answer (1.11).** When  $E$  and  $F$  are independent, finding  $P(E \cap F)$  is straightforward because the sample space is simply the set of all pairs of possible outcomes from  $E$  and  $F$ , i.e.,  $\Omega = \Omega_E \times \Omega_F$ , where  $\Omega_G$  is the sample space of an event  $G$ . As a result,

$$P(E \cap F) = \frac{\#(E \cap F)}{|\Omega_E \times \Omega_F|} = \frac{\#(E) \cdot \#(F)}{|\Omega_E| \cdot |\Omega_F|} = P(E) \cdot P(F),$$

where  $\#(G)$  denotes the number of outcomes where event  $G$  occurs (in its sample space). See Figure A.1 for an example.

	1	2	3	4	5
4	(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)
3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)
2	(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)
1	(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)

Figure A.1: The elements of  $\Omega_{D_5} \times \Omega_{D_4}$ . Let  $X$  be the event where  $D_5$  is prime, and let  $Y$  be the event where  $D_4$  is even. The outcomes where  $X \cap Y$  occurs are boxed, and we can verify that  $P(X \cap Y) = 6/20$  matches  $P(X) \cdot P(Y) = \frac{3}{5} \cdot \frac{2}{4}$ .

**Answer (1.12).** Venn diagram properties shown in Figure A.2 reveal that for any two events  $E$  and  $F$ ,

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

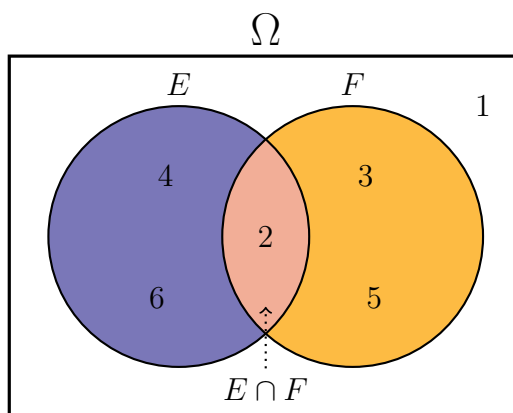


Figure A.2: A Venn diagram of two events in d6's sample space.  $E$  is the event that the roll is even, and  $F$  is the event that the roll is prime. So  $P(E \cup F) = P(E) + P(F) - P(E \cap F) = 3/6 + 3/6 - 1/6 = 5/6$ .

**Answer (1.13).** We first notice that

$$(x + h)^2 = x^2 + 2hx + h^2$$

so

$$\left(x + \frac{b}{2a}\right)^2 = x^2 + \frac{b}{a}x + \frac{b^2}{4a^2}.$$

Then

$$\begin{aligned}ax^2 + bx + c &= 0 \\x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} + \frac{c}{a} &= \frac{b^2}{4a^2} \\ \left(x + \frac{b}{2a}\right)^2 + \frac{c}{a} &= \frac{b^2}{4a^2} \\ \left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \\ x + \frac{b}{2a} &= \pm \sqrt{\frac{b^2}{4a^2}} \\ x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.\end{aligned}$$