The Algebra of Binding

Distinction, Interaction, and Admissible Collapse
from Elementary Identities to General Systems

Flyxion

2026

Preface: Why Cross-Terms Matter

Take three lengths, x, y, and z, and build a square whose side is their sum. The area of that square is $(x+y+z)^2$. Divide the square along the three lengths and nine regions appear: three along the diagonal, with areas x^2, y^2, and z^2, and six off the diagonal, with areas xy, xz, and yz, each appearing twice. The full expansion is

$$(x+y+z)^2 = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz.$$

Every student who has expanded a trinomial square has produced this identity. Few are asked to say where the six rectangles come from. The usual answer is procedural: the distributive law requires them, FOIL generalized to nine terms requires them, and that is the end of it. This book asks a different question. What do the rectangles *mean*, and what do they mean about systems that are not squares of side lengths at all?

The diagonal terms, x^2, y^2, z^2, are what each side contributes on its own. They would exist even if the square were never assembled, as long as x, y, and z were separately measured. The off-diagonal terms exist only because the three sides were placed together into one figure. They are not properties of x, or of y, or of z individually. They are properties of the arrangement.

This is the distinction the book is built on. Call the sum of the diagonal terms the *separate evaluation*:

$$Q_{\text{separate}} = Q(x) + Q(y) + Q(z).$$

Call the evaluation of the assembled whole the *bound evaluation*:

$$Q_{\text{bound}} = Q(x+y+z).$$

For the square, these differ by exactly the six rectangles:

$$\Delta_B = Q_{\text{bound}} - Q_{\text{separate}} = 2xy + 2xz + 2yz.$$

Call Δ_B the *binding correction*. It is the amount by which joining components changes what can be said about them, relative to what could be said about each in isolation.

The claim of this book is that Δ_B is worth taking seriously well beyond geometry. Variance of a sum of random variables differs from the sum of variances by exactly a covariance term of the same algebraic shape. Portfolio risk differs from the sum of individual asset risks by exactly the same shape of term. Mutual information, mutual coupling in a graph, the difference between evaluating a database table alone and evaluating it after a join: all of these are instances, in a suitably general sense, of the same pattern. Something is true of the whole that is not the sum of what is true of the parts, and the discrepancy has a name, a source, and in many cases an exact formula.

The book will insist on a distinction that the elementary case makes easy to state and later cases make easy to forget. There are really three separate claims bundled together whenever someone says a whole is more than the sum of its parts. The first is a theorem: for a specific Q and a specific decomposition, Δ_B has a computable value, possibly zero. The second is a piece of mathematics one level up: this holds not only for squares of sums but for a whole class of quadratic and bilinear evaluations, and the class can be characterized precisely. The third is a modeling choice: applying this structure to an epistemic, computational, ecological, or social system requires someone to say, explicitly, what plays the role of x, y, z, what plays the role of Q, and what operation plays the role of "assembling into one figure." The first two claims are provable. The third is not automatic, and treating it as if it inherited the certainty of the first two is the most common way this kind of reasoning goes wrong.

So a caution belongs here, stated as plainly as the thesis. The area model makes every cross-term positive, because lengths are nonnegative and area is additive over disjoint regions. Nothing forces this outside geometry. In a general bilinear pairing $B(x, y)$, the correction may be positive, negative, or zero: constructive in some couplings, destructive in others, absent where two components are independent or orthogonal with respect to the evaluation in question. A theory that only knows how to say "the whole exceeds the sum of the parts" has thrown away half of what the algebra actually offers. The rigorous content of this book is not that binding produces surplus. It is that binding produces a *term*, of a specific algebraic shape, whose sign and magnitude must be derived or measured rather than assumed.

One more distinction, easy to miss in the collected form of the identity, will recur throughout the book. Written out before collection, the nine products of $(x+y+z)(x+y+z)$ include both xy and yx as separate entries in the multiplication table, arising from different orders of selection. Ordinary commutativity lets us write $xy + yx = 2xy$, and nothing is lost numerically. But something is lost representationally: the collected form no longer shows that two distinguishable routes produced that coefficient. Most of the systems this book applies the framework to are not commutative in the relevant sense, and the difference between a history that recorded two routes and a state that recorded one number becomes exactly the question of whether the system can be audited, repaired, or run backward. The elementary identity is a good place to see this cheaply, because in ordinary arithmetic the loss costs nothing. Later it will cost quite a lot.

The book is organized to move outward from this single figure in five stages: from the identity itself to the general algebra of quadratic and bilinear forms; from that algebra to a small calculus of four operations, distinguishing, refusing, binding, and collapsing, that describe how systems are built up and simplified; from the calculus to a sequence of substantive domains, probability, graphs, dynamics, computation, evidence, value, where the same binding correction appears under different names; and finally to the question that motivates the whole project, which is not whether a total can be computed, but whether the way it was computed preserves what a later reader, auditor, or repair process will need. The closing chapters argue that this last question, admissibility of collapse, is the same question in every domain the book visits, asked in different vocabularies.

The six rectangles, in other words, are not a detour on the way to the answer. In this book they are treated as the more interesting part of the picture than the answer itself.

Chapter 1

Multiplication as Pair Formation

1.1 The question behind FOIL

Most students learn to expand $(a+b)(c+d)$ by a mnemonic: First, Outer, Inner, Last.

$$(a+b)(c+d) = \underbrace{ac}_{\text{First}} + \underbrace{ad}_{\text{Outer}} + \underbrace{bc}_{\text{Inner}} + \underbrace{bd}_{\text{Last}}.$$

The mnemonic is a memory aid for a sequence of four multiplications, and it is usually retired once the pattern is automatic. This chapter retires the mnemonic in the other direction: instead of forgetting it once it is no longer needed, it asks what operation FOIL was a special case of, so that the operation, rather than the four-letter word, is what survives into the rest of the book.

The answer is exhaustive pairing. Given two factors, each a sum of terms, multiplying them means forming every ordered pair consisting of one term from the first factor and one term from the second, multiplying each pair, and adding the results. FOIL names four such pairs because a binomial has two terms and $2 \times 2 = 4$. Nothing in the underlying operation refers to the number two. It generalizes immediately to any number of terms in either factor, and the four-letter mnemonic simply stops having enough letters.

1.2 Indexed sums and the distributive law

Let $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ be two finite sequences of terms, and consider the product

$$\left(\sum_{i=1}^{m} a_i \right) \left(\sum_{j=1}^{n} b_j \right).$$

The distributive law, applied repeatedly, gives

$$\left(\sum_{i=1}^{m} a_i \right) \left(\sum_{j=1}^{n} b_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j.$$

The right-hand side is a sum over the Cartesian product of index sets $\{1, \ldots, m\} \times \{1, \ldots, n\}$. Each pair (i, j) in that product contributes exactly one term, $a_i b_j$, to the sum. There are mn such pairs, and correspondingly mn terms before anything is simplified.

For $m = n = 2$, this is FOIL:

$$(a_1 + a_2)(b_1 + b_2) = a_1b_1 + a_1b_2 + a_2b_1 + a_2b_2,$$

with $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$ the four ordered pairs, named First, Outer, Inner, Last by their position in the written expression rather than by anything intrinsic to the pairs themselves. The names are typographical. The pairs are combinatorial. This distinction matters more than it looks like it should, because later in the book terms will be relabeled, reordered, and rewritten in bases where "first" and "last" no longer mean anything, while the underlying pairing structure persists.

Definition 1.1. *Given finite sequences $(a_i)_{i\in I}$ and $(b_j)_{j\in J}$, the* expanded product *is the indexed family*

$$\big(a_ib_j\big)_{(i,j)\in I\times J},$$

and the product value *is $\sum_{(i,j)\in I\times J} a_ib_j$.*

The expanded product and the product value are different objects. The expanded product is a family indexed by $I \times J$: it remembers, for every term, which pair of source terms produced it. The product value is a single number or expression, obtained by adding up the family. Two different expanded products can have the same product value. This is the first appearance of a distinction the book will not let go of: a *generative history* (here, the indexed family) and a *derived value* (here, the sum), which need not determine each other in reverse.

1.3 The case of equal factors

The special case that this book cares about most is $a_i = b_i$ for all i: squaring a sum rather than multiplying two different sums.

$$\left(\sum_{i=1}^{n} x_i\right)^2 = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=1}^{n} x_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_ix_j.$$

Here the index set for both factors is the same, $\{1,\dots,n\}$, and the pairs (i,j) range over $\{1,\dots,n\}^2$, an $n \times n$ grid rather than a general $m \times n$ rectangle. This grid splits naturally into two parts:

$$\{1,\dots,n\}^2 = \{(i,i) : 1 \le i \le n\} \ \cup \ \{(i,j) : i \ne j\}.$$

The first part, the diagonal, contains n pairs. The second part, the off-diagonal, contains $n^2 - n$ pairs, which split further into $\binom{n}{2}$ unordered pairs $\{i,j\}$ with $i < j$, each realized twice, once as (i,j) and once as (j,i):

$$n^2 = \underbrace{n}_{\text{diagonal}} + \underbrace{2\binom{n}{2}}_{\text{off-diagonal}}.$$

This is an identity about counting, true before any variable is assigned a value and before any commutativity is invoked. It says only that the $n \times n$ grid of ordered pairs decomposes into n pairs with equal entries and $2\binom{n}{2}$ pairs with distinct entries, arranged in $\binom{n}{2}$ mirrored couples across the diagonal.

Example 1.2. *For $n = 3$, indices $\{1, 2, 3\}$, the nine pairs are*

$$(1,1), (2,2), (3,3) \quad \textit{(diagonal, three pairs)}$$

$$(1,2), (2,1), \quad (1,3), (3,1), \quad (2,3), (3,2) \quad \textit{(off-diagonal, six pairs in three mirrored couples).}$$

Writing $x_1 = x$, $x_2 = y$, $x_3 = z$ recovers the multiplication table from the preface, and the count $9 = 3 + 2\binom{3}{2} = 3 + 6$ matches it exactly.

1.4 Where commutativity enters, and where it does not

Nothing said so far has used the fact that ordinary multiplication commutes. The pairing structure, the grid of ordered pairs (i, j), and the diagonal/off-diagonal split are true for exhaustive pairing over any two index sets, regardless of what algebraic properties the terms being multiplied happen to have. Commutativity enters only at the next step, when the two entries of a mirrored couple, $x_i x_j$ and $x_j x_i$, are asserted to be equal:

$$x_i x_j = x_j x_i \quad \Longrightarrow \quad x_i x_j + x_j x_i = 2 x_i x_j.$$

This step is doing two things at once, and the rest of the book depends on keeping them apart. It is asserting a numerical fact, that the two products have the same value, and it is licensing a representational move, collecting two entries of the indexed family into a single scalar multiple. The numerical fact is true whenever the underlying multiplication commutes: real numbers, complex numbers, ordinary lengths and areas. The representational move, collapsing an indexed pair into a coefficient, is available whenever the numerical fact holds, but taking it does not follow automatically from the fact being available. One could, in principle, always keep $x_i x_j$ and $x_j x_i$ as separate entries in a ledger even though they are numerically equal, if the ledger's purpose required remembering that there were two of them and where they came from.

This chapter will not yet say why anyone would want to do that. That question belongs to Chapter 3. What this chapter establishes is only the precondition for asking it: that behind every collected coefficient, such as the 2 in $2xy$, there is an exhaustive pairing structure that produced a specific number of specific ordered pairs, and that this structure exists and is countable independently of whether anything is later done to simplify it.

1.5 Summary

Multiplication of two sums is exhaustive pairing: every term of the first factor is paired with every term of the second, each pair contributes one product, and the pairs are indexed by the Cartesian product of the two index sets. FOIL is the special case of two terms in each factor, and the names First, Outer, Inner, Last describe typographical position, not mathematical structure. Squaring a sum of n terms produces an $n \times n$ grid of pairs, splitting into n diagonal pairs and $\binom{n}{2}$ off-diagonal mirrored couples, a purely combinatorial fact established before commutativity is invoked. Commutativity is what allows a mirrored couple to be collected into a single coefficient; it is a separate step from the pairing itself, and the next chapter turns to what that collected coefficient means when the terms being combined are not simply the sides of a geometric figure.

Exercises

1. Write out the expanded product for $(a_1 + a_2 + a_3)(b_1 + b_2)$ as an indexed family over $\{1, 2, 3\} \times \{1, 2\}$, and confirm it has six terms.

2. For $n = 4$, compute the number of diagonal and off-diagonal pairs in $\{1, \ldots, 4\}^2$ directly by listing them, and check the count against $n + 2\binom{n}{2}$.

3. Construct an example of two finite sequences (a_i) and (b_j), drawn from a non-commutative setting (for instance, 2×2 matrices), where $a_i b_j \neq b_j a_i$ in general. Explain which parts of this chapter's argument still hold and which no longer do.

4. Explain, in your own words, why "the expanded product" and "the product value" are different mathematical objects even when they determine the same number in a given case.

Chapter 2

The Binding Correction

2.1 Additivity, and its failure

A function L is additive if

$$L(x+y) = L(x) + L(y)$$

for all x and y in its domain. Linear maps have this property by definition, and a great deal of elementary mathematics is built on the convenience of additive quantities: lengths along a line, masses of disjoint objects, probabilities of disjoint events. When a quantity is additive, evaluating the parts and adding the results gives the same answer as evaluating the whole. Nothing is lost by decomposing.

The square function is not additive. For real numbers,

$$(x+y)^2 = x^2 + 2xy + y^2 \neq x^2 + y^2$$

whenever $xy \neq 0$. Chapter 1 showed where the extra term comes from: it is the mirrored couple $xy + yx$, collected into $2xy$, arising from the off-diagonal pairs that exist only because x and y were multiplied against each other rather than evaluated separately. This chapter takes that single example and turns it into a general statement about any evaluation that behaves this way, whether or not it happens to be literally a square.

Definition 2.1. *Let Q be a function on a set closed under an addition operation. Define the* binding correction *of Q at x and y as*

$$\Delta_Q(x,y) = Q(x+y) - Q(x) - Q(y).$$

Δ_Q measures exactly the failure of additivity, no more and no less. If Q is additive, $\Delta_Q \equiv 0$. If Q is not additive, Δ_Q is itself a well-defined function of the pair (x,y), computable once Q is specified, and it is this function, not the mere observation that $Q(x+y) \neq Q(x) + Q(y)$, that the rest of the book treats as the object of interest.

Example 2.2. *For $Q(x) = x^2$,*

$$\Delta_Q(x,y) = (x+y)^2 - x^2 - y^2 = 2xy.$$

For $Q(x) = x^3$,

$$\Delta_Q(x,y) = (x+y)^3 - x^3 - y^3 = 3x^2y + 3xy^2.$$

For $Q(x) = e^x$ under ordinary addition,

$$\Delta_Q(x, y) = e^{x+y} - e^x - e^y,$$

which does not simplify to a small polynomial in x and y, but is still a perfectly well-defined function, computable for any x, y, and vanishing nowhere on the reals except along one curve.

The third case is worth pausing on. Nothing in the definition of Δ_Q requires Q to be a polynomial, or the correction to have a clean closed form. What makes the quadratic case ($Q(x) = x^2$) the organizing example for this book is not that it is the only case, but that it is the simplest case in which Δ_Q is itself bilinear, meaning it is linear in x for fixed y and linear in y for fixed x. That property is special, and Part II is devoted to it. This chapter deals with Δ_Q in general, before specializing.

2.2 Many terms

For n components $x_1, \ldots, x_n$, the definition extends directly. Write

$$\Delta_Q(x_1, \ldots, x_n) = Q\left(\sum_{i=1}^{n} x_i\right) - \sum_{i=1}^{n} Q(x_i).$$

For the square,

$$Q\left(\sum_i x_i\right) = \sum_i x_i^2 + 2\sum_{i<j} x_i x_j,$$

so

$$\Delta_Q(x_1, \ldots, x_n) = 2\sum_{i<j} x_i x_j.$$

This is the same computation as the trinomial expansion in the Preface, now stated for arbitrary n rather than fixed at three. The diagonal terms, $\sum_i x_i^2$, reproduce the separate evaluation $\sum_i Q(x_i)$ exactly, by construction, since $Q(x_i) = x_i^2$. Every other term in the expansion belongs to Δ_Q. This is not a coincidence particular to squaring: for any Q, the separate evaluation $\sum_i Q(x_i)$ is defined to be whatever is left after the binding correction is subtracted off, so the decomposition

$$Q\left(\sum_i x_i\right) = \sum_i Q(x_i) + \Delta_Q(x_1, \ldots, x_n)$$

holds by definition for any Q whatsoever. What is not automatic, and what has to be established case by case, is whether Δ_Q has a usable form: a formula, a sign, a decomposition into pairwise and higher-order pieces, or a physical or evidential interpretation. For the square, Δ_Q decomposes entirely into pairwise terms, with no three-way or higher contributions. This is a genuine fact about $Q(x) = x^2$, not a fact about binding corrections in general; Part XI (Chapter 18 in this condensed sequence) returns to evaluations whose corrections do not decompose so simply.

2.3 What the correction does and does not tell you

It is tempting, once Δ_Q has a name, to treat any nonzero value of it as evidence that something interesting happened when x and y were combined: interaction, coupling, emergence, synergy. This temptation should be resisted, and resisting it is the main substantive claim of this chapter.

$\Delta_Q(x, y)$ is a fact about the function Q applied to the specific values x and y. It is not, by itself, a fact about a process, a mechanism, or a relationship between whatever x and y represent. Three qualifications follow.

First, the sign and magnitude of Δ_Q depend entirely on the choice of Q. For $Q(x) = x^2$, $\Delta_Q(x, y) = 2xy$ is positive whenever x and y have the same sign, negative whenever they have opposite signs, and zero whenever either is zero. None of this reflects a judgment about whether combining x and y was good, bad, or consequential; it reflects only the algebra of squaring. Choosing a different Q, on the same x and y, produces a different correction with a different sign pattern. The correction is a property of the pairing (Q, x, y), not of x and y alone.

Second, a correction of zero does not establish that x and y are unrelated, only that this particular Q does not register their relation. $\Delta_Q(x, y) = 0$ can arise from genuine independence, from a coincidental cancellation, from a Q that is simply insensitive to the kind of relation that holds between x and y, or from measurement at a resolution too coarse to see it. Chapter 5 returns to this point at length, under the heading of orthogonality, once the machinery for distinguishing these cases has been built.

Third, and most importantly for how this book intends to be used, applying Δ_Q outside of a setting where x, y, and Q are all precisely specified numbers is not yet a mathematical claim. If x and y are meant to stand for two people's competencies, two species' population sizes, or two pieces of evidence, then someone has to say what Q is, in enough detail that $Q(x+y)$, $Q(x)$, and $Q(y)$ can actually be computed or measured, before Δ_Q means anything beyond a suggestive analogy. The algebra in this chapter is exact and general. Its application to a substrate where x and y are not numbers is a separate, additional piece of work, and the applications chapters later in the book (Part IV) do that work explicitly, one substrate at a time, rather than asserting it by resemblance to the formula.

Proposition 2.3. *Let Q be additive on its domain. Then $\Delta_Q(x, y) = 0$ for all x, y in that domain.*

Proof. Immediate from the definitions: $\Delta_Q(x, y) = Q(x + y) - Q(x) - Q(y) = 0$ is exactly the additivity condition. □

The converse of Proposition 2.3 is false in general: Δ_Q can vanish at a specific pair (x, y) without Q being additive everywhere. For $Q(x) = x^2$, $\Delta_Q(x, y) = 2xy = 0$ whenever $x = 0$ or $y = 0$, even though Q is additive nowhere else. A single vanishing correction is a local fact about a pair of values, not a global fact about the function.

2.4 The correction as the object of study

The reframing this chapter proposes is small but consequential. Rather than asking whether $Q(x + y)$ equals $Q(x) + Q(y)$, a yes-or-no question with no further content once answered, the book asks what $\Delta_Q(x, y)$ is: a specific, computable quantity, varying with x, y, and the choice of Q, capable of being positive, negative, or zero, and available for further analysis in its own right rather than as a leftover.

This is the same move already made once, in Chapter 1, when the collected coefficient $2xy$ was unpacked into two separate entries xy and yx from the expanded product. There, the question was where a number came from. Here, the question is what a discrepancy is made of. Both moves refuse to let a single scalar stand in for the structure that produced it. The next chapter asks what is lost, and what can still be recovered, when that refusal is not maintained: when two different generative histories are allowed to collapse into the same total, and the total alone is kept.

Exercises

1. Compute $\Delta_Q(x, y)$ for $Q(x) = |x|$ (absolute value) and determine the set of pairs (x, y) for which it vanishes.

2. Compute $\Delta_Q(x, y)$ for $Q(x) = \max(x, 0)$ and describe its sign pattern across the four sign-combinations of x and y.

3. For $Q(x) = x^2$ and three terms x, y, z, verify by direct expansion that $\Delta_Q(x, y, z) = 2(xy + xz + yz)$, and confirm this matches $\Delta_Q(x, y) + \Delta_Q(x, z) + \Delta_Q(y, z)$, i.e., that the three-term correction is exactly the sum of the pairwise corrections. State, without proving it yet, whether you expect this additivity of corrections to hold for $Q(x) = x^3$.

4. Give an example, drawn from any domain you like, where two quantities x and y plausibly interact in some causal or physical sense, but a poorly chosen Q would report $\Delta_Q(x, y) = 0$ anyway. Identify what a better-chosen Q would need to be sensitive to.

Chapter 3

Provenance and Collection of Like Terms

3.1 Two objects, one number

Chapter 1 distinguished the expanded product, an indexed family recording which pair of source terms produced each entry, from the product value, the single number or expression obtained by summing that family. Chapter 2 showed that the discrepancy between a bound evaluation and a separate evaluation is itself a well-defined quantity, Δ_Q, deserving study in its own right rather than dismissal as a correction term. This chapter asks a question that sits underneath both: when a generative process is summarized by its total, what exactly has been kept, and what has been given up.

Return to the trinomial expansion. Before anything is simplified,

$$(x + y + z)(x + y + z) = x^2 + xy + xz + yx + y^2 + yz + zx + zy + z^2,$$

nine terms, each traceable to one ordered pair of indices. Ordinary commutativity licenses $xy = yx$, and the expression is usually written

$$x^2 + y^2 + z^2 + 2xy + 2xz + 2yz,$$

six terms, three of them coefficients attached to a single symbol standing for two. Both expressions denote the same number for any real x, y, z. They are not the same expression.

Definition 3.1. *Let H denote a* generative history*: an indexed family of terms together with the operations that produced them, retaining the distinctness of terms that are numerically equal but arose from different generative steps. Let $V(H)$ denote the* derived value*: the number or simplified expression obtained by evaluating and combining the entries of H under whatever equivalences (commutativity, associativity, arithmetic identity) apply to the terms involved.*

The nine-term expansion is a generative history. The six-term expansion is one possible derived value of it, obtained by applying the equivalence $xy \sim yx$ and collecting equal entries. A further, more aggressive derived value is available too: if $x = 1, y = 2, z = 3$, then $V(H) = 36$, a single number with no visible trace of squares or cross-terms at all. Each step, from history to six-term expression to bare number, is a legitimate simplification. Each step also discards something the previous representation still contained.

3.2 The general fact

Proposition 3.2. *Let H_1 and H_2 be two generative histories, possibly arising from different processes, different orderings, or different numbers of steps. It is possible for $V(H_1) = V(H_2)$ while $H_1 \neq H_2$.*

Proof. By example. Let H_1 be the history that produces xy once and yx once, as two separate entries, and let H_2 be the history that produces $2xy$ directly, as a single entry, by some process that never generates yx at all (for instance, a rule that doubles a single computed product rather than computing the product twice in two orders). For real x, y, $V(H_1) = xy+yx = 2xy = V(H_2)$, yet H_1 has two entries and a specific pairing structure that H_2 never had. □

The proof is almost too simple to feel like it is proving anything, and that is the point of placing it here rather than reserving this style of argument for a setting where it looks more serious. The claim

$$V(H_1) = V(H_2) \implies H_1 = H_2$$

is true for the same structural reason in every setting the rest of this book visits: a merge log that reconstructs a file and a from-scratch write that produces an identical file; two independent measurements that happen to agree and a single measurement copied twice; two witnesses who corroborate each other and one witness quoted twice under different names. The algebra is not being used metaphorically in these later cases. It is being used to certify, in each case, exactly what the arithmetic already certifies here: that agreement in value is compatible with, and gives no information about, agreement in history. Whether the specific applications of that fact are apt is a question for the chapters that make them (Chapter 12, on corroboration and dependence, is the first). What is not in question is the underlying algebraic fact itself.

3.3 When collection is not free

Chapter 1 flagged, as an exercise, a setting where the collection step is not merely representationally costly but numerically illegitimate: multiplication that does not commute. It is worth working through explicitly here, because it draws a sharp line between two very different reasons one might refuse to write $2xy$ in place of $xy + yx$.

Example 3.3. *Let A and B be 2×2 matrices,*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

so $AB \neq BA$, and consequently $AB + BA \neq 2AB$. Any expression that wrote $2AB$ in place of $AB + BA$ would simply be wrong, not merely less detailed.

This gives two distinct failure modes for a collected expression, and they should not be run together. In the commutative case, real numbers, ordinary lengths, $xy = yx$ is true, $2xy$ is numerically correct, and what is lost by writing it is provenance: the record of two distinguishable generative steps that happened to agree in value. Retaining $xy + yx$ instead of $2xy$ is a choice

about what the representation is for, not a correction to an error. In the non-commutative case, matrices, quaternions, most operations on states and processes, $AB \neq BA$ in general, and $2AB$ is not an available derived value of the history that produced AB once and BA once; it denotes something else, computed a different way, that happens to coincide with the true total only in special cases (for instance, whenever A and B commute with each other specifically, even if matrix multiplication in general does not). The first failure mode is a loss of information under a valid equivalence. The second is the use of an equivalence that does not hold.

Most of this book's applications live closer to the second case than the first, once they are examined carefully. Two pieces of evidence bearing on a hypothesis, two edits applied to a shared document, two agents' actions in a shared environment: whether their joint effect is order-independent is a substantive question about the domain, not something to be assumed by analogy with ordinary multiplication. Chapter 6 (Directed Bindings, in the condensed sequence) returns to this distinction under the heading of directionality, once the vocabulary of bilinear forms is available to state it precisely.

3.4 What a history-preserving representation requires

If $V(H_1) = V(H_2)$ carries no information about whether $H_1 = H_2$, then any task that depends on knowing which history actually occurred, auditing a computation, repairing a corrupted record, adjudicating a disputed claim, cannot be discharged by the derived value alone. It requires access to H itself, or to enough of H to answer the question at hand.

This does not mean derived values are dispensable or that collection is always a mistake. Most uses of arithmetic want exactly the derived value and nothing more: a total area, a final balance, a distance traveled. The claim is narrower and more specific: that collection is a choice with a cost, that the cost is zero for some purposes and prohibitive for others, and that the purpose has to be specified before the choice can be evaluated. A book on binding and interaction that only ever computed totals would have nothing further to say once Chapter 2 was finished. What makes the rest of this book possible is that some of its central questions, about audit, repair, evidence, and admissibility, are exactly the questions for which the derived value is not enough, and for which the discarded history in Proposition 3.2 has to be recovered, retained, or reasoned about directly.

3.5 Closing Part I

The three chapters of Part I form a single argument. Multiplication of sums is exhaustive pairing, producing an indexed family of terms before anything is simplified (Chapter 1). When the evaluation in question is not additive, the discrepancy between its value on a sum and the sum of its values on the parts is itself a computable, analyzable quantity, not merely a sign that something has gone wrong (Chapter 2). And the step from an indexed family of terms to a single collected value is a genuine step, licensed under some equivalences and not others, discarding provenance even when it is numerically sound (Chapter 3).

Part II takes up the case where the evaluation Q is specifically quadratic, meaning Δ_Q is itself bilinear, and asks what extra structure that assumption buys. The trinomial square has been the working example throughout Part I precisely because it is quadratic in this sense; Chapter 4 makes that assumption explicit and general, replacing the fixed example with the full algebra of quadratic and bilinear forms.

Exercises

1. Construct two different generative histories $H_1 \neq H_2$, distinct from the one given in this chapter, such that $V(H_1) = V(H_2)$ under ordinary real-number arithmetic.

2. For the matrix example, compute $AB - BA$ and confirm it is nonzero. This quantity, the commutator of A and B, will reappear in later chapters as one measure of how strongly an order dependence matters.

3. Give an everyday (non-mathematical) example of two different histories producing the same summary, and identify precisely what information the summary discards.

4. Suppose a system's designer chooses to store only $V(H)$, never H itself, in order to save space. Describe one task this choice makes impossible, and one task it does not affect at all. What distinguishes the two tasks?

Part I

Quadratic Forms and Polarization

Chapter 4

Quadratic Forms and Bilinear Corrections

4.1 Why quadratic

Part I worked with one running example, $Q(x) = x^2$, and one general definition, $\Delta_Q(x, y) = Q(x+y)-Q(x)-Q(y)$, applicable to any Q whatsoever. The generality was deliberate: additivity, and its failure, are meaningful questions for absolute value, for the exponential, for any function at all. But the example was not chosen at random, and this chapter finally says why.

For $Q(x) = x^2$, the correction $\Delta_Q(x, y) = 2xy$ has a property that Δ_Q for $|x|$ or e^x does not: it is linear in x when y is held fixed, and linear in y when x is held fixed. Doubling x doubles the correction; adding two values of x before computing the correction gives the same result as computing two corrections and adding them. This property is called bilinearity, and its presence or absence is what separates the quadratic case, in the technical sense this chapter defines, from the merely nonlinear case. Quadratic evaluations are worth a dedicated Part of this book because their corrections are bilinear, and bilinear corrections are the ones for which the richest and most exact theory is available: a matrix representation, a notion of orthogonality, a change-of-basis theory, and, in Chapter 5, a formula for recovering the correction itself from evaluations of Q alone.

4.2 Quadratic forms

Definition 4.1. *A function $Q : \mathbb{R}^n \to \mathbb{R}$ is a* quadratic form *if it can be written as*

$$Q(\mathbf{x}) = \mathbf{x}^\mathsf{T} A \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

for some $n \times n$ matrix $A = (a_{ij})$.

Without loss of generality A may be taken to be symmetric ($a_{ij} = a_{ji}$), since the antisymmetric part of any matrix contributes zero to $\mathbf{x}^\mathsf{T} A \mathbf{x}$: if $a_{ij} \neq a_{ji}$, replacing both entries with their average $(a_{ij} + a_{ji})/2$ leaves Q unchanged. This book follows the standard convention and assumes A symmetric from here on, unless a specific application requires otherwise (Chapter 6 discusses when it does).

Example 4.2. *The trinomial square is the quadratic form with $n = 3$ and A the all-ones matrix,*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \qquad Q(x, y, z) = (x, y, z)A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x + y + z)^2.$$

Every entry of A is 1: each variable contributes its full square to the diagonal, and every pair of distinct variables is coupled with the same strength. This is the most symmetric possible quadratic form on three variables, and it is precisely this uniformity that made the trinomial square a clean example in Part I. Once A is allowed to vary, that uniformity becomes a special case rather than a default.

Expanding the double sum and separating diagonal from off-diagonal terms, exactly as in Chapter 1,

$$Q(\mathbf{x}) = \sum_i a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j.$$

The diagonal entries a_{ii} govern each variable's self-contribution; the off-diagonal entries a_{ij} govern the strength of the pairwise coupling between x_i and x_j. Setting every $a_{ii} = 1$ and every $a_{ij} = 1$ recovers the trinomial square. Setting $a_{ij} = 0$ for $i \neq j$ and $a_{ii} = 1$ recovers the fully separate evaluation $\sum_i x_i^2$, with no binding correction at all. Every quadratic form on n variables lies somewhere between these two extremes, and the matrix A is exactly the data needed to say where.

4.3 Bilinear forms

Definition 4.3. *A function $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a* bilinear form *if it is linear in each argument separately:*

$$B(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = B(\mathbf{x}_1, \mathbf{y}) + B(\mathbf{x}_2, \mathbf{y}), \qquad B(c\mathbf{x}, \mathbf{y}) = cB(\mathbf{x}, \mathbf{y}),$$

and symmetrically in the second argument.

Every bilinear form on $\mathbb{R}^n$ can be written $B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\mathsf{T} A \mathbf{y}$ for some matrix A, and conversely every such expression defines a bilinear form. The connection to quadratic forms is immediate: given a symmetric bilinear form B, setting $Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x})$ produces a quadratic form, and every quadratic form arises this way from its associated symmetric matrix.

Proposition 4.4. *Let $Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x})$ for a symmetric bilinear form B. Then*

$$\Delta_Q(\mathbf{x}, \mathbf{y}) = Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y}) = 2B(\mathbf{x}, \mathbf{y}).$$

Proof.

$$Q(\mathbf{x} + \mathbf{y}) = B(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = B(\mathbf{x}, \mathbf{x}) + B(\mathbf{x}, \mathbf{y}) + B(\mathbf{y}, \mathbf{x}) + B(\mathbf{y}, \mathbf{y}),$$

using bilinearity to expand. By symmetry of B, $B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x})$, so

$$Q(\mathbf{x} + \mathbf{y}) = Q(\mathbf{x}) + Q(\mathbf{y}) + 2B(\mathbf{x}, \mathbf{y}),$$

and subtracting $Q(\mathbf{x}) + Q(\mathbf{y})$ from both sides gives the claim. □

This is the exact general form of the computation carried out by hand in the Preface and in Chapter 2. It says something Part I could only illustrate one instance at a time: for *every* quadratic form, not merely $Q(x) = x^2$, the binding correction is precisely twice the corresponding bilinear form evaluated on the two arguments. The bilinear form B is not a separate piece of data one must additionally specify. It is already implicit in Q, recoverable from it, a fact Chapter 5 turns into a working method.

4.4 Reading off the matrix

Because $\Delta_Q(\mathbf{x}, \mathbf{y}) = 2\mathbf{x}^\mathsf{T} A\mathbf{y}$ for the same matrix A that defines Q, the off-diagonal entries of A are directly interpretable as pairwise coupling strengths, in the following precise sense: writing $\mathbf{e}_i$ for the i-th standard basis vector,

$$\Delta_Q(\mathbf{e}_i, \mathbf{e}_j) = 2\mathbf{e}_i^\mathsf{T} A\mathbf{e}_j = 2a_{ij} \qquad (i \neq j).$$

The binding correction between the i-th and j-th coordinate directions, taken alone, recovers exactly twice the (i, j) entry of A. This gives a second, equivalent way to think about what a quadratic form's matrix encodes: not merely a table of coefficients to be multiplied out, but a table of measured or specified pairwise binding corrections, one for each pair of coordinates, together with diagonal self-contributions of a genuinely different kind. The diagonal case is worth working out explicitly rather than assumed by analogy: $\Delta_Q(\mathbf{e}_i, \mathbf{e}_i) = Q(2\mathbf{e}_i) - 2Q(\mathbf{e}_i) = 4a_{ii} - 2a_{ii} = 2a_{ii}$, which does *not* vanish, unlike the naive expectation that a self-correction ought to be zero by symmetry with additivity. The resolution is that a_{ii} and $\Delta_Q(\mathbf{e}_i, \mathbf{e}_i)$ answer two different questions: a_{ii} is the coefficient of the self-term x_i^2 in Q, the diagonal entry itself, while $\Delta_Q(\mathbf{e}_i, \mathbf{e}_i) = 2a_{ii}$ is what the correction *formula* returns when applied to a coordinate direction and itself, a quantity with no privileged interpretive role here and not one this book will call a "self-correction" going forward, precisely to avoid suggesting it should vanish. The entries a_{ii} remain what this book means by self-contributions; the off-diagonal a_{ij}, recovered via $\Delta_Q(\mathbf{e}_i, \mathbf{e}_j) = 2a_{ij}$ for $i \neq j$, remain what it means by pairwise binding corrections. This is exactly the diagonal/off-diagonal split from Chapter 1, now named on both sides.

4.5 What this buys

Restating the trinomial square as a quadratic form with a fully symmetric all-ones matrix makes visible a question Part I had no vocabulary to ask: what happens to the theory when the couplings are not uniform, when some pairs of variables are strongly bound and others are not bound at all? The answer is that nothing in Proposition 4.4 required uniformity. Any symmetric A defines a valid quadratic form, with its own bilinear correction, its own pattern of pairwise couplings, and its own diagonal self-terms, independent of whether that pattern happens to be uniform, sparse, or anything in between. Chapter 9 (Sparse Binding, in the condensed sequence) develops the sparse case directly, once graphs have been introduced as the natural language for saying which pairs are coupled at all.

The next chapter takes up a different question: given only the ability to evaluate Q itself, at arbitrary points, without ever being told the matrix A or the form B, can the coupling be recovered? The answer, called polarization, turns Proposition 4.4 around: instead of computing Δ_Q from a known B, it computes B from measured values of Q alone.

Exercises

1. Write the quadratic form for $Q(x, y) = 3x^2 + 5y^2 + 4xy$ in matrix form, identifying A explicitly, and compute $\Delta_Q(x, y)$ directly from the definition. Confirm it matches $2\mathbf{x}^\mathsf{T} A\mathbf{y}$ when $\mathbf{x} = (x, 0)$ and $\mathbf{y} = (0, y)$.

2. For the matrix A with $a_{11} = a_{22} = 1$ and $a_{12} = a_{21} = 0$, write out $Q(x, y)$ and confirm $\Delta_Q(x, y) = 0$ for all x, y. What does this say about the geometric picture from the Preface if the two rectangles were removed from the figure entirely?

3. Construct a symmetric 2×2 matrix A with $a_{12} < 0$, and describe, using Proposition 4.4, the circumstances under which $\Delta_Q(x, y)$ is negative despite x and y both being positive.

4. For general n, show that the number of independent entries in a symmetric $n \times n$ matrix A (counting a_{ij} and a_{ji} as one entry) is $n + \binom{n}{2}$, and relate this count to the diagonal/off-diagonal split from Chapter 1.

Chapter 5

Polarization and Orthogonality

5.1 Turning the formula around

Proposition 4.4 computes Δ_Q from a known bilinear form B: given B, the correction is $2B(\mathbf{x}, \mathbf{y})$. This chapter runs the computation in the other direction. Suppose only Q is known, or only measurable, and B is not given in advance. Since $\Delta_Q(\mathbf{x}, \mathbf{y}) = 2B(\mathbf{x}, \mathbf{y})$ and Δ_Q is defined entirely in terms of Q, rearranging gives

$$B(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}\big[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})\big].$$

This identity is called *polarization*. It says that the bilinear form associated with a quadratic form is not extra data one must be separately handed. It is already determined by Q, recoverable from three evaluations of Q alone: at $\mathbf{x} + \mathbf{y}$, at $\mathbf{x}$, and at $\mathbf{y}$.

Proposition 5.1 (Polarization identity)**.** *Let $Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x})$ for a symmetric bilinear form B. Then for all $\mathbf{x}, \mathbf{y}$,*

$$B(\mathbf{x}, \mathbf{y}) = \tfrac{1}{2}\big[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})\big].$$

Proof. Immediate from Proposition 4.4, dividing both sides by 2. □

The identity has a companion form, obtained by expanding $Q(\mathbf{x} - \mathbf{y})$ the same way and combining:

$$B(\mathbf{x}, \mathbf{y}) = \tfrac{1}{4}\big[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x} - \mathbf{y})\big].$$

Both forms recover the same B; the first uses three evaluations of Q, the second uses two. Which is more convenient depends on the setting: the second is standard in physics, where $Q(\mathbf{x} + \mathbf{y})$ and $Q(\mathbf{x} - \mathbf{y})$ might correspond to two directly measurable configurations (constructive and destructive combination), while the first is closer to the bookkeeping already used in this book, where $Q(\mathbf{x})$ and $Q(\mathbf{y})$ are the separate evaluations from Chapter 2 and $Q(\mathbf{x} + \mathbf{y})$ is the bound evaluation.

5.2 What polarization is a method for

Polarization is best understood not as a formula to memorize but as a method: a way of discovering relational structure by comparing isolated and joint behavior, without ever inspecting the internal composition of $\mathbf{x}$ or $\mathbf{y}$ more closely. If Q is the only thing that can be measured (a total

energy, a total risk, a total variance) and B (the pairwise coupling) is what is actually wanted, polarization says the coupling is already implicit in three numbers: the joint measurement and the two isolated measurements. Nothing about the internal structure of $\mathbf{x}$ or $\mathbf{y}$ needs to be known separately from what Q reports.

Example 5.2. *Suppose a quadratic form Q is known only through evaluations, and the values $Q(1,0) = 3$, $Q(0,1) = 5$, $Q(1,1) = 12$ have been measured (in the notation of the previous chapter's first exercise, this is exactly $Q(x,y) = 3x^2 + 5y^2 + 4xy$, though nothing here assumes that formula is known in advance). Polarization gives*

$$B\big((1,0),(0,1)\big) = \tfrac{1}{2}\big[12 - 3 - 5\big] = 2,$$

recovering the off-diagonal entry $a_{12} = 2$ exactly, from three measurements, without ever needing to see the matrix A directly. (Note that $B(\mathbf{e}_1, \mathbf{e}_2) = a_{12}$ is not itself the full cross-term coefficient in Q; the cross-term in $Q(x,y) = 3x^2 + 5y^2 + 4xy$ is $4xy = 2a_{12}xy$, consistent with Proposition 4.4.) The two-evaluation form gives the same answer by a different route: $Q(1,-1) = 3 + 5 - 4 = 4$, so $B\big((1,0),(0,1)\big) = \frac{1}{4}[12 - 4] = 2$.

This is the precise sense in which the binding correction, introduced informally in Chapter 2 as "a quantity worth studying in its own right," turns out for quadratic forms to be a fully determined and fully recoverable piece of structure: not a residual left over after the interesting work is done, but the interesting work itself, extractable from nothing more than the evaluation function.

5.3 Orthogonality

Definition 5.3. *Two vectors $\mathbf{x}$ and $\mathbf{y}$ are* orthogonal with respect to Q *(or B-orthogonal) if $B(\mathbf{x}, \mathbf{y}) = 0$.*

By Proposition 4.4, $B(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\Delta_Q(\mathbf{x}, \mathbf{y}) = 0$, that is, if and only if $Q(\mathbf{x}+\mathbf{y}) = Q(\mathbf{x}) + Q(\mathbf{y})$. Additivity, which Chapter 2 treated as a property that either holds globally for a function or does not, reappears here as something that can hold *locally*, at a specific pair of vectors, for a function that is not additive in general. For the form with matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, every pair of distinct standard basis vectors is orthogonal, since $a_{12} = 0$, even though Q itself, $Q(x,y) = x^2 + y^2$, is not an additive function of a general vector. For the trinomial-square matrix (all entries 1), by contrast, no two nonzero coordinate vectors are orthogonal, since every $a_{ij} = 1 \neq 0$.

This sharpens a warning first raised in Chapter 2. There, it was noted that $\Delta_Q(x,y) = 0$ does not by itself establish that x and y are unrelated, only that this particular Q does not register their relation. The vocabulary of orthogonality makes the qualification exact rather than merely cautionary: $\Delta_Q(\mathbf{x}, \mathbf{y}) = 0$ establishes precisely that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal *with respect to this particular B*, nothing more and nothing less. Whether that is evidence of independence, cancellation, or simply an unlucky choice of Q depends entirely on what B is supposed to represent, a question external to the algebra.

Example 5.4. *Let $Q(\mathbf{x}) = \operatorname{Var}\left(\sum_i x_i X_i\right)$ for random variables $X_1, \ldots, X_n$ and coefficients x_i (this anticipates Chapter 10, where the connection between variance and quadratic forms is*

developed fully). The associated bilinear form is $B(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}\left(\sum_i x_i X_i, \sum_j y_j X_j\right)$*, and B-orthogonality of two coordinate directions* $\mathbf{e}_i, \mathbf{e}_j$ *is exactly* $\mathrm{Cov}(X_i, X_j) = 0$*: uncorrelatedness. It is a standard and important fact, taken up again in Chapter 11, that uncorrelated random variables need not be independent. Orthogonality with respect to the variance quadratic form is a strictly weaker condition than probabilistic independence, and the gap between them is not a defect of the algebra; it is exact information about how much a covariance measurement can and cannot tell you.*

5.4 Absent cross-terms have more than one cause

Given a measured or observed $\Delta_Q(\mathbf{x}, \mathbf{y}) = 0$, at least four distinct explanations are available, and polarization gives no way to distinguish them from the number zero alone.

The two quantities may be genuinely independent in whatever sense matters for the application, with B correctly reflecting that independence. The two quantities may be dependent in a way this specific B is not built to detect, as in Example 5.4: a nonlinear dependence produces zero covariance routinely. Two nonzero relational contributions may cancel exactly, as when $B(\mathbf{x}, \mathbf{y})$ sums several distinct coupling mechanisms with opposite sign; the zero total does not mean no coupling occurred, only that whatever occurred summed to zero. And the measurement or model of Q may simply be too coarse, aggregating away a real relation before B is ever computed, in which case zero reflects a limitation of the measurement rather than a fact about the underlying relation at all.

None of these four cases is distinguishable from any other by the value $\Delta_Q(\mathbf{x}, \mathbf{y}) = 0$ on its own. Distinguishing them requires either a finer Q, a different Q altogether, or information from outside the quadratic-form framework entirely, such as an intervention that manipulates one quantity and observes whether the other responds (Chapter 19, on intervention on relations, develops this route directly). The lesson of this section is not that vanishing correlation is useless information; it is real information, and often exactly the information wanted. The lesson is that it is one specific, limited kind of information, and treating a zero as a general certificate of "no relation" overstates what the algebra has actually shown.

5.5 Where this leaves the theory, and where it does not

Polarization completes the picture opened in Chapter 4. A quadratic form is fully equivalent to its bilinear form; the two determine each other exactly, in both directions, and the binding correction between any two vectors is available from evaluations of Q alone, with no separate access to B required. Orthogonality gives a precise, local notion of "no correction here," sharper than the global all-or-nothing additivity from Chapter 2. But precision about what a vanishing correction *is* (a fact about B at a specific pair) is not the same as an account of what it *means* outside the algebra, and Example 5.4 together with the four-fold list above are meant to keep that gap visible rather than let the cleanliness of the polarization formula paper over it.

One further question remains before Part II is finished. Everything so far has been stated relative to a fixed choice of coordinates, the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, and the matrix A has been treated as though it were an intrinsic description of the coupling among $x_1, \ldots, x_n$. It is not quite that. The next chapter asks what happens to A, to the diagonal/off-diagonal split, and to the very appearance of cross-terms at all, when the coordinates themselves are changed.

Exercises

1. For $Q(x, y) = 3x^2 + 5y^2 + 4xy$, compute $B\big((1, 1), (1, -1)\big)$ using polarization, and confirm it agrees with a direct computation from the matrix $A = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}$ via $B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathsf{T}} A \mathbf{y}$.

2. Construct a quadratic form on $\mathbb{R}^2$ for which $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal but $\mathbf{e}_1$ and $\mathbf{e}_1 + \mathbf{e}_2$ are not. What does this show about whether orthogonality is a property of a pair of coordinate directions or a property of the individual vectors involved?

3. Give an example (numerical or descriptive) of two quantities with a strong nonlinear relationship but zero covariance, distinct from the standard example of a symmetric parabola. Identify explicitly which of the four explanations in this chapter's list applies.

4. Suppose $\Delta_Q(\mathbf{x}, \mathbf{y}) = 0$ is observed, and an intervention that changes $\mathbf{x}$ while holding the mechanism generating $\mathbf{y}$ fixed causes Δ_Q to become nonzero. Which of the four explanations does this rule out, and which remain possible?

Chapter 6

Basis Dependence

6.1 A vanishing act

Every computation in the last two chapters was carried out in one fixed coordinate system, the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, in which $x_1, \ldots, x_n$ are the separately addressable components whose self-terms and cross-terms make up Q. This chapter asks what happens to that whole apparatus when the coordinates are changed.

Take the trinomial-square matrix once more, A with every entry equal to 1, so $Q(x, y, z) = (x + y + z)^2$. Introduce a new coordinate,

$$w = \frac{x + y + z}{\sqrt{3}},$$

together with any two further coordinates completing w to an orthonormal basis of $\mathbb{R}^3$ (for instance, $u = (x - y)/\sqrt{2}$ and $v = (x + y - 2z)/\sqrt{6}$; the specific choice will not matter). A direct computation, carried out in the exercises, shows

$$Q = (x + y + z)^2 = 3w^2 + 0 \cdot u^2 + 0 \cdot v^2.$$

In the new coordinates (w, u, v), the quadratic form has *no cross-terms at all.* Every off-diagonal entry of the matrix representing Q in this basis is zero. The six rectangles that organized the entire Preface, the pairwise couplings that Chapters 4 and 5 spent two chapters characterizing, are simply absent from this description.

Nothing dishonest has happened. Q is the same function; it assigns the same number to the same physical point regardless of which coordinates are used to name that point. The apparent disappearance of the cross-terms is entirely a fact about the coordinates, not about Q.

6.2 Where the relation went

The temptation, on seeing $Q = 3w^2$, is to conclude that the interaction between x, y, and z was never real, that it was an artifact of a poorly chosen basis, and that the "true" description of the system has no binding correction whatsoever. This conclusion should be resisted, but it should be resisted for a more precise reason than the informal one that follows immediately from it. The full change of coordinates,

$$\mathbf{w} = P^{\mathsf{T}}\mathbf{x}, \qquad \mathbf{x} = P\mathbf{w},$$

is invertible: P is orthogonal, so $P^{-1} = P^{\mathsf{T}}$, and $\mathbf{x}$ can be recovered from $\mathbf{w}$ exactly, for any point, by applying P. This means that the triple (w, u, v), *together with the matrix P that relates it back to* (x, y, z), contains exactly as much information as (x, y, z) did. Nothing has been discarded. Any question answerable from (x, y, z), including "how much did x contribute," remains answerable, by first recovering $\mathbf{x} = P\mathbf{w}$ and then asking the question in the original coordinates. An invertible re-description is not a loss of information; it is a relabeling, and relabelings do not, by themselves, discard anything.

What is true, and what motivated the informal statement above, is narrower and more specific. The coordinate $w = (x + y + z)/\sqrt{3}$ is a specific linear combination of all three original variables, exactly the combination the cross-terms were coupling together, and the self-term $3w^2$ reproduces the total contribution of that combination in a single number. Asking "how much does w alone contribute" is a different question from "how much does x alone contribute," and the second question, if it is to be answered from (w, u, v) rather than from (x, y, z), requires converting back: computing $\mathbf{x} = P\mathbf{w}$ first, and only then isolating x. The information needed to do this, the matrix P, is not itself part of the quadratic form Q; it is separate data describing which coordinate change was used. If P is retained alongside (w, u, v), the original components remain fully recoverable, and no relation has been lost, only relocated into a form ($3w^2$ plus P) that requires an extra step to unpack. If P is discarded, or if only w is kept while u and v are dropped because Q happens not to depend on them, then something genuinely has been lost, and it is exactly the ability to answer "how much did x contribute" that goes with it.

This is the precise sense in which basis change and information loss are not the same operation, and conflating them would understate what an invertible change of coordinates actually preserves. Diagonalizing a quadratic form, on its own, relocates relational information from explicit cross-terms into the definition of new coordinates and the transformation relating them back to the old ones; it does not, by itself, discard that information. What can discard it is a further step: throwing away P, or projecting (w, u, v) down to w alone on the grounds that Q does not need u and v. That further step, not the basis change itself, is where a genuine loss occurs, and Chapter 8 will have reason to treat it as a specific case of Collapse precisely because it is a projection, not because it is a change of coordinates.

6.3 Eigenvalues and the general picture

The trinomial-square example is a special case of a completely general fact.

Theorem 6.1 (Spectral theorem, stated without proof)**.** *Every symmetric matrix A can be written $A = PDP^{\mathsf{T}}$, where P is orthogonal (its columns are an orthonormal basis of eigenvectors of A) and D is diagonal (its entries are the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ of A).*

In the coordinates given by the columns of P, any quadratic form $Q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}} A\mathbf{x}$ becomes

$$Q = \sum_i \lambda_i w_i^2,$$

with no cross-terms whatsoever, where $\mathbf{w} = P^{\mathsf{T}}\mathbf{x}$. This holds for *every* symmetric A, not merely the all-ones matrix: every quadratic form, however densely coupled its original variables appear to be, can be re-expressed with zero off-diagonal entries in a suitably chosen basis. For the trinomial square, A has eigenvalues $3, 0, 0$ (the reader is asked to verify this in the exercises), which is why only one term, $3w^2$, survives: the other two eigendirections carry zero

self-contribution because Q, restricted to those directions, is identically zero. A matrix of full rank with all-distinct nonzero eigenvalues would diagonalize to n nonzero self-terms rather than one, but would still have zero cross-terms in its eigenbasis.

This generality sharpens the point of the previous section rather than undermining it. If *every* coupling can be diagonalized away by a suitable choice of coordinates, then the presence or absence of cross-terms in a given representation of Q alone cannot, by itself, be read as a fact about whether the underlying components are related, as long as the full coordinate change and its inverse are kept in view. What cannot be changed by any choice of basis are the eigenvalues themselves, $3, 0, 0$ in the running example: these are basis-independent invariants of Q, unlike the individual entries a_{ij} of A, which depend entirely on which basis was chosen to write A down. The number of nonzero eigenvalues (the *rank* of Q) says something real and coordinate-free about the form: for the trinomial square, rank 1, meaning Q itself, however many cross-terms it displays in the standard basis, is really only sensitive to a single combined quantity, the total $x + y + z$. This is a fact about Q, and it is compatible with x, y, z remaining fully individually recoverable from any complete coordinate description; Q's insensitivity to u and v does not imply that u and v themselves, or the components they are built from, have become inaccessible.

6.4 Admissible basis change

The question this chapter has been circling is not whether a basis change is mathematically legal. Any invertible linear change of coordinates is legal, and the spectral theorem guarantees a particularly convenient one always exists. The question is what a basis change, or more precisely what is done *after* a basis change, is required to preserve in order to remain admissible for a given purpose, a question that has no single universal answer and must be posed relative to that purpose.

At minimum, an admissible basis change preserves the value of Q itself: $Q(\mathbf{x})$ computed from the original coordinates and $Q(P^{\mathsf{T}}\mathbf{x})$ computed from the new ones must agree, and orthogonal changes of basis guarantee this automatically. If the full change of coordinates is retained, together with P, considerably more is preserved as well: the ability to recover $\mathbf{x}$ exactly, and with it the ability to answer any question the original coordinates could answer, including questions about x, y, or z individually. The loss this chapter has been circling arises only from a further, separable decision: to keep Q's value but discard the means of recovering the original components, whether by discarding P outright or by discarding the eigencoordinates (u, v in the running example) that Q happens not to depend on. That decision, not the basis change, is what forecloses the question "how much did x contribute," and it is worth naming as a distinct step because it is possible to change basis without ever taking it.

6.5 Closing Part II

Part II has given the trinomial square its full general treatment. Any quadratic evaluation is equivalent to a symmetric matrix (Chapter 4); the pairwise couplings recorded in that matrix are recoverable from evaluations of Q alone, via polarization, and a vanishing coupling is a precise but limited local fact rather than a general certificate of independence (Chapter 5); and the couplings visible in a given matrix are an artifact of the chosen coordinates, in the sense that every quadratic form can be re-expressed with no couplings at all in its eigenbasis, while

the information needed to recover the original components remains fully available provided the coordinate change itself is not thrown away (this chapter).

That final qualification supplies the transition into Part III. An invertible change of coordinates, on its own, is not the kind of operation that discards information; only a subsequent projection, keeping Q's value while dropping the means of recovering the components it was built from, does that. Distinguishing these two operations precisely, one that merely re-describes a system and one that genuinely narrows what can later be asked of it, is exactly the job Part III takes up. It introduces four operations by name, Pop, Refuse, Bind, and Collapse, and it will treat invertible re-description as something other than Collapse: Collapse, in the sense Part III defines, is reserved for the operation that actually identifies distinct histories or discards distinct options, which the eigenbasis change, taken together with P, does not do. Projecting away P or the null coordinates u, v, keeping only w, is a Collapse in exactly this sense, and it is that projection, not the diagonalization itself, that the next Part's admissibility apparatus is built to evaluate.

Exercises

1. Verify by direct substitution that $(x+y+z)^2 = 3w^2$ where $w = (x+y+z)/\sqrt{3}$, $u = (x-y)/\sqrt{2}$, $v = (x+y-2z)/\sqrt{6}$, by expressing x, y, z in terms of w, u, v and expanding. (It is enough to confirm the coefficient of u^2 and v^2 vanishes and cross-terms among w, u, v vanish; the full inverse transformation is not required for this exercise, but is requested in Exercise 4.)

2. Confirm that the all-ones 3×3 matrix has eigenvalues $3, 0, 0$, by verifying that $(1,1,1)$ is an eigenvector with eigenvalue 3 and that any vector orthogonal to $(1,1,1)$ is an eigenvector with eigenvalue 0.

3. Take the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Find its eigenvalues and eigenvectors, write $Q(x,y) = 2x^2 + 2y^2 + 2xy$ in the eigenbasis, and confirm the off-diagonal term vanishes there.

4. Write down the explicit inverse transformation P taking (w, u, v) back to (x, y, z) for the trinomial-square example. Using it, show that the contribution of x alone (for instance, the value of x at a specific point, or x^2 evaluated there) is fully recoverable from (w, u, v) together with P, even though it cannot be read off directly from the single term $3w^2$. Then describe what would need to be discarded, P itself, or the coordinates u, v, for that recoverability to actually fail.

Part II

A Calculus of Distinction and Binding

Chapter 7

Worlds, Histories, and Option Spaces

7.1 What the algebra has been assuming

Parts I and II worked entirely with values: numbers, vectors, matrices, quadratic forms. Even where the argument turned on something other than a value, the generative history H of Chapter 3, the eigenbasis of Chapter 6, that something was still fixed once and for all at the start of the discussion, given rather than constructed. This chapter begins the work of treating construction itself as the object of study: not just what Q evaluates to, but the sequence of operations by which the objects being evaluated came to be distinguished, related, and simplified in the first place.

The motivation is the loose end left at the close of Chapter 6. Choosing a basis, diagonalizing a matrix, replacing x, y, z with w, u, v: these were described there as things one *does* to a quadratic form, with consequences for what remains recoverable afterward. But "things one does" has not yet been given any formal status. This chapter supplies the setting in which it can be: a notion of a system's authoritative history, a notion of what remains open given that history, and a precise relationship between the two that mirrors, and generalizes, the history-versus-value distinction from Chapter 3.

7.2 Worlds

Definition 7.1. *A* world *is a pair $W = (H, \Omega)$, where H is an* authoritative history*, a sequence of events or operations that have occurred, and Ω is an* option space*, the set of continuations, extensions, or alternative operations that remain available given H.*

H is a record, not a summary: it is a sequence, with an order, in the same sense that the nine-term expanded product of Chapter 1 was a sequence of nine distinguishable ordered pairs before anything was collected. Ω is not a prediction of what will happen, nor a probability distribution over futures; it is the space of what is *admissible* to happen next, given everything recorded in H so far.

H and Ω play different formal roles, but they are not independent: the world's governing semantics fixes a map

$$\omega : \mathcal{H} \to \mathcal{O}, \qquad \Omega = \omega(H),$$

determining, for any history, exactly which continuations are admissible from it. H records what occurred; Ω records which continuations that occurrence leaves open; and Ω is a function

of H under whatever admissibility rules the world in question is governed by, not a second independent input one is free to specify separately. Writing $W = (H, \Omega)$ as a pair, rather than simply working with H and computing $\omega(H)$ whenever Ω is needed, is a representational choice rather than a claim of independence: it lets Ω be tracked and consulted directly, without recomputing it from the complete history each time, which matters once H becomes long or its early portions become costly to reprocess. This is best understood as a state carried alongside H subject to the invariant $\Omega = \omega(H)$, and any operation defined on worlds in this book is required to preserve that invariant: an operation that changes H must update the recorded Ω to match ω of the new history, and an operation that appears to change Ω without a corresponding change to H is not well-formed under this framework.

A world with a rich H typically has a more constrained Ω than a world with a sparse H: the more that has been committed, the fewer the admissible continuations. This asymmetry, history closing off options rather than merely accumulating facts, is the reason Ω is worth tracking explicitly rather than only ever computing on demand; two histories of the same apparent length can leave very different option spaces behind, depending on what kind of commitments they contain, and the tracked Ω makes that difference immediately visible without requiring ω to be recomputed from scratch.

7.3 States are derived, not substituted

Definition 7.2. *A* state *is a value* $S = \text{state}(H)$*, computed from a history* H *by some fixed rule* $\text{state}(\cdot)$*, that summarizes* H *for some specific purpose.*

This is exactly the relationship between H and $V(H)$ from Chapter 3, restated in the vocabulary of worlds. A state is a derived value; a history is the generative record it was derived from. Proposition 3.2 transfers immediately: two different histories, $H_1 \neq H_2$, can produce the same state, $\text{state}(H_1) = \text{state}(H_2)$, and this is not a defect to be engineered away but an ordinary and expected feature of any summarizing rule worth using. The point of insisting on the vocabulary of worlds, rather than continuing to talk about histories and values directly, is to make explicit a consequence of this fact that Chapter 3 did not need to dwell on: if a system is described *only* by its current state S, with H discarded once S has been computed, then the system's future option space Ω cannot in general be recovered from S alone.

Example 7.3. *Let* H *record a sequence of deposits and withdrawals to an account, and let* $S = \text{state}(H)$ *be the resulting balance. Two different histories, one consisting of a single large deposit, another consisting of many small deposits and a few withdrawals that happen to net to the same total, can produce identical balances* S*. If the account is subject to a rule such as "no more than three withdrawals per statement period," the option space* Ω*, what withdrawals remain admissible right now, depends on which withdrawals have already occurred in* H*, not merely on the current balance. Two accounts with identical balances can have entirely different admissible next moves.*

This is not a special fact about banking. It is the general shape of every case in this book where a summary was shown to discard something a later operation needs: the eigenbasis of Chapter 6 discarded the ability to attribute an outcome to x alone; the eigenbasis is, in the vocabulary of this chapter, a state, and the original coordinates together with the operation that produced the eigenbasis are the history from which that state was derived. Whether the discarded information matters depends entirely on what the option space Ω needs to contain

going forward, which is a claim about the world's purpose, not about the state's arithmetic correctness.

7.4 Distinctions as elements of a world

The vocabulary of Parts I and II can now be restated inside a single world. A distinction, an x, a y, a coordinate direction, an addressable component, is something that has been separately entered into H: it exists as an identifiable entry in the history, with its own position, prior to any operation that might later relate it to another distinction or collapse it together with one. This is deliberately the same notion of "distinguishable" at work in Chapter 1's indexed pairing, where x_i and x_j were addressable separately before their product $x_i x_j$ was formed, and in Chapter 3's generative history, where xy and yx were separate entries before collection. A world's history H is, among other things, a record of which distinctions have been made so far and in what order.

Given this, the four operations that Chapter 8 will define formally, distinguishing (creating a new addressable entry), refusing (recording that two entries must not be treated as equivalent), binding (coupling two entries so that their joint evaluation carries a relational term), and collapsing (deriving a state from the history, in the sense of this chapter, thereby closing off some part of Ω), are not being introduced as new machinery unrelated to the algebra already developed. They are names for operations this book has been performing informally since the Preface, made explicit so that their effects on Ω, and not merely their effects on a value $V(H)$ or state(H), can be tracked.

7.5 Option-Space Sufficiency

Exercise 4, in an earlier draft of this chapter, was left as an open question: does an injective state map σ, one that never identifies two different histories, guarantee that nothing has been lost? The honest answer is that injectivity is sufficient but not necessary, and that the actual condition of interest is weaker and more informative than injectivity. This section states it directly, rather than leaving it as an exercise, because the distinction it draws is one the rest of Part III depends on.

Let $\mathcal{H}$ denote the set of histories under consideration, $\mathcal{S}$ the set of possible states, and $\mathcal{O}$ the set of possible option spaces. Write

$$\sigma : \mathcal{H} \to \mathcal{S}$$

for the state map, $\sigma(H) = \text{state}(H)$, and

$$\omega : \mathcal{H} \to \mathcal{O}$$

for the map sending each history to the option space it leaves behind.

Proposition 7.4 (Option-space sufficiency)**.** *The state $\sigma(H)$ preserves all information required to determine the current option space if and only if ω is constant on every fiber of σ, that is,*

$$\sigma(H_1) = \sigma(H_2) \implies \omega(H_1) = \omega(H_2)$$

for all $H_1, H_2 \in \mathcal{H}$. Equivalently, ω factors through σ: there exists a map $\bar{\omega} : \mathcal{S} \to \mathcal{O}$ with $\omega = \bar{\omega} \circ \sigma$.

Proof. If ω factors through σ, then $\sigma(H_1) = \sigma(H_2)$ immediately gives $\omega(H_1) = \bar{\omega}(\sigma(H_1)) = \bar{\omega}(\sigma(H_2)) = \omega(H_2)$. Conversely, if ω is constant on every fiber $\sigma^{-1}(S)$ for $S \in \sigma(\mathcal{H})$, define $\bar{\omega}(S)$ to be that constant value; this is well-defined precisely because the fiber is not split by ω, and by construction $\bar{\omega} \circ \sigma = \omega$. □

Call a state map satisfying this condition *option-sufficient*. Injectivity of σ is a special case: if σ never identifies two distinct histories, its fibers each contain a single history, and ω is trivially constant on a one-element set. But a state map can identify many histories, discarding a great deal of provenance in the sense of Chapter 3, and still be option-sufficient, provided every history merged into a common state happens to leave the same option space behind. The account-balance example from earlier in this chapter illustrates the failure mode precisely: a rule such as "no more than three withdrawals per period" makes ω sensitive to how many withdrawals have occurred, a fact the balance alone does not determine, so $\sigma(H_1) = \sigma(H_2)$ can hold while $\omega(H_1) \neq \omega(H_2)$. That balance rule is option-insufficient with respect to that particular constraint on Ω. A simpler account, with no such constraint, would have the balance as an option-sufficient (indeed usually the natural and complete) state, despite still merging infinitely many distinct deposit-and-withdrawal histories into each balance value.

The danger condition this chapter has actually been describing, then, is not non-injectivity as such. It is

$$\exists\, H_1, H_2 \in \mathcal{H} : \quad \sigma(H_1) = \sigma(H_2) \;\text{ but }\; \omega(H_1) \neq \omega(H_2),$$

two histories collapsed to the same state despite genuinely different continuations being admissible from them. This is the precise, checkable form of what later chapters will call *premature collapse*: a summary is premature, relative to a task, exactly when it merges histories that the task still needs to tell apart.

7.6 Three cases, and two kinds of recoverability

This gives three cases rather than the earlier two, and it is worth naming them, since Part III will refer back to this distinction repeatedly. An *injective* state preserves the entire history, and therefore trivially preserves every history-determined option space, along with everything else the history contains. An *option-sufficient but non-injective* state loses historical detail, in the sense of Chapter 3, provenance, order, the specific path taken, without losing anything the current option space depends on: the loss is real but harmless relative to ω, though it may not be harmless relative to some other purpose (an audit, for instance, that cares about the path itself rather than merely what remains admissible). An *option-insufficient* state collapses histories whose available continuations genuinely differ, and is premature with respect to any task that needs Ω correctly determined.

One further distinction is worth making explicit before moving on, since it is easy to conflate with the three cases above. Injectivity of σ guarantees that H is recoverable from S *in principle*: a unique history corresponds to each state, so an inverse function exists. It does not guarantee that inverse is recoverable *in practice*. If computing $\sigma^{-1}(S)$ requires resources unavailable to whoever needs the answer, unbounded search, an intractable computation, access to auxiliary records that were not retained, the history has not been informationally destroyed, but it has been rendered effectively irrecoverable under the relevant resource bound. The two failures look identical from the outside, a question about H that cannot be answered from S, but they call

for different remedies: the first requires retaining more of H; the second requires retaining, or being able to reconstruct, the means of inverting σ.

Recoverability, in short, is not a single question. Before asking whether a given state or collapse "preserves enough," the specific target of recovery has to be named: the full history, the present option space, a future update to that option space, or something else again. Chapter 8 will use option-space sufficiency, Proposition 7.4, as its working admissibility criterion for the Collapse operation specifically, precisely because it is the version of recoverability that connects most directly to what a world can still do next, rather than to what it has already done.

7.7 Why this is worth the added formalism

It is fair to ask whether a book that has so far gotten real mileage out of ordinary algebra, quadratic forms, matrices, eigenvalues, needs a new vocabulary of worlds, histories, and option spaces at all. The honest answer is that the algebra alone cannot express the claim this chapter has been building toward: that two representations can agree on every value computed so far and still differ in what they permit next. Nothing in the notation $Q(\mathbf{x})$, Δ_Q, or $B(\mathbf{x}, \mathbf{y})$ has a place to record an option space, because a quadratic form is a function from vectors to numbers, full stop; it has no notion of what remains open. The moment this book's questions turn from "what is the value" to "what can still happen, and what has been foreclosed," as they did informally already in the discussion of provenance in Chapter 3 and admissibility in Chapter 6, the vocabulary of values alone runs out, and something like (H, Ω) becomes necessary rather than decorative.

Part III accordingly shifts register, from evaluating fixed algebraic objects to describing operations on worlds. The formulas of Parts I and II do not disappear; they reappear inside Part III and afterward as one particular kind of state$(\cdot)$, the kind that computes a scalar value from a history of distinctions and bindings. What changes is that this book can now also ask, of any such computation, what it cost: which entries of H it is no longer possible to recover, and which elements of Ω it has quietly removed.

Exercises

1. Using the account-history example, construct two specific histories H_1 and H_2 with state$(H_1) =$ state(H_2) but different admissible option spaces $\Omega_1 \neq \Omega_2$ under the three-withdrawals rule. Confirm this shows the balance map is option-insufficient with respect to that rule.

2. Recast the trinomial-square construction of the Preface as a world: what does H contain at the point where x, y, z have been introduced but not yet squared together, and what does H contain after $(x + y + z)^2$ has been computed? Is anything in Ω different between the two points?

3. Construct a state map that is non-injective (it merges at least two distinct histories) but is nonetheless option-sufficient in the sense of Proposition 7.4. Identify the fiber that gets merged and explain why ω agrees on it.

4. Give an example of a state map that is injective but computationally impractical to invert under a stated resource bound (for instance, a one-way permutation $\sigma : \{0,1\}^n \to \{0,1\}^n$, bijective by construction, for which computing σ is efficient but computing σ^{-1} is believed

to require infeasible search; note that this differs from a fixed-length cryptographic hash, which maps a larger domain to a smaller range and therefore cannot be injective at all, only collision-resistant). Explain why the permutation example differs from the account-balance example, where the loss is informational rather than merely practical.

Chapter 8

Pop, Refuse, Bind, Collapse

8.1 Four operations on a world

Chapter 7 introduced a world $W = (H, \Omega)$ and argued, via Proposition 7.4, that the question of what a derived state preserves has a precise answer: a state $\sigma(H)$ is option-sufficient exactly when it does not merge histories whose option spaces differ. This chapter names four specific operations that act on worlds, each changing H, Ω, or both, and states the admissibility question for each of them in terms of that same criterion.

Definition 8.1. *For a history H, let $E(H)$ denote the set of addressable entries currently popped in H: the entries introduced by* Pop *operations recorded in H so far.*

Definition 8.2. *Given a world $W = (H, \Omega)$, define four operations.*

- Pop(a)*: extends H by introducing a new addressable entry $a \notin E(H)$, and updates $\Omega = \omega(H)$ accordingly to reflect that a is now available to be referenced, refused, or bound.*
- Refuse(a, b)*, for $a, b \in E(H)$: extends H by recording the pair (a, b) as forbidden to identify, and updates $\Omega = \omega(H)$ accordingly, removing from the current admissible continuation set any continuation that would apply a state map identifying a and b.*
- Bind(a, b)*, for $a, b \in E(H)$: extends H by recording a coupling between entries a and b, and updates $\Omega = \omega(H)$ accordingly to reflect that subsequent evaluations may depend on their joint configuration rather than on either alone.*
- Collapse$_\sigma(H)$*: applies a state map σ to the current history, producing $S = \sigma(H)$, and updates $\Omega = \omega(H)$ accordingly to whatever continuations remain determinable from S alone.*

Each operation is defined so as to preserve the invariant of Chapter 7, $\Omega = \omega(H)$: extending H and then recomputing ω is the only way Ω ever changes under this framework.

Three of these four operations, Pop, Refuse, and Bind, only ever add to H; they are, in the vocabulary of Chapter 7, purely constructive, and H itself is never edited or shrunk by any of the four, only extended. The fourth, Collapse, is the one operation that deliberately trades access to H for a smaller, more manageable object S, and it is therefore the only one of the four for which the question of admissibility, in the sense developed in Chapter 7, actually arises.

Pop, Refuse, and Bind can fail to be well-formed (one cannot bind an entry that has not been popped, for instance, which is why both are required to draw their arguments from $E(H)$), but they cannot be premature in Chapter 7's sense, because they never collapse anything. This asymmetry is the organizing fact of the present chapter.

8.2 Pop

Pop(a) is the operation that makes a new distinction addressable. Every variable in Parts I and II, every x_i in an indexed sum, entered the discussion this way: before x and y could be multiplied, added, or compared, each had to be separately available as something with its own identity. The chapter is careful to separate three things that Pop does not, by itself, accomplish. Popping a creates the entry; it does not yet assign a a numerical value, since a distinction can be addressable while still symbolic, as every variable in this book has been prior to any specific substitution. It does not yet relate a to anything else, since Bind is a separate operation. And it does not by itself guarantee a is distinct in the sense of Refuse: two entries can be popped separately and later found, or declared, to be equivalent, and the machinery for recording that decision belongs to the next section, not to Pop itself.

8.3 Refuse

Refuse(a, b), for entries $a, b \in E(H)$, records that a and b are not to be identified. This is a constraint on future Collapse operations, not an evaluation: it says nothing about the numerical relationship between a and b (they might later turn out to have the same value under some σ), only that no admissible continuation may treat them as the same entry. Formally, let

$$R_H \subseteq E(H) \times E(H)$$

denote the set of entry pairs recorded as refused at some point in H; note the type is pairs of *entries*, not pairs of histories, since a refusal is always a claim about two specific addressable distinctions, not about two entire alternative histories.

Checking whether a proposed Collapse$_\sigma$ respects R_H requires knowing what it would mean for σ to "identify" two entries, and this is not automatic: σ is a map out of whole histories, $\sigma : \mathcal{H} \to \mathcal{S}$, and does not by itself act on individual entries at all. Making Refuse-compliance checkable therefore requires σ to be paired with an induced relation

$$\sim_\sigma \subseteq E(H) \times E(H),$$

declaring which entries σ's derived state treats as equivalent; this is additional structure that must be supplied alongside σ, not something every state map carries automatically. A subsequent Collapse$_\sigma$ is inadmissible with respect to R_H if

$$\exists\, (a, b) \in R_H \quad \text{such that} \quad a \sim_\sigma b.$$

If no such induced relation has been specified for a given σ, whether it respects R_H is simply not yet a well-posed question, and specifying $\sim_\sigma$ is part of what it means to fully specify σ for use in a world where refusals have been recorded.

One further point is worth making explicit, since it is easy to state loosely in a way that looks self-contradictory. A refusal $(a, b) \in R_H$ is never removed from H: H is append-only, and

the historical fact that a and b were refused persists in the record permanently, exactly like any other entry. What changes is Ω, the currently admissible continuation set, computed as $\omega(H)$: once $(a, b) \in R_H$, any continuation applying a σ with $a \sim_\sigma b$ is no longer a member of $\omega(H)$. "Recorded permanently in H" and "removed from the current Ω" describe two different objects, the append-only history and the derived, and potentially shrinking, option space, and are not in tension: the first never shrinks, the second can, and Refuse's entire effect is on the second, mediated by the first.

Refuse is therefore best understood as pre-emptively shrinking the set of state maps that may legitimately be applied later, rather than as an operation with content of its own independent of some eventual Collapse.

Example 8.3. *Suppose two witnesses' reports, a and b, are entered into a case history separately, so $a, b \in E(H)$. If it later emerges that both reports trace back to the same original source, so that treating them as independent corroboration would be an error (an application taken up in full in Chapter 12), a* Refuse(a, b) *operation, or its substantive analogue, records that any later scoring rule σ whose induced relation $\sim_\sigma$ treats a and b as interchangeable corroboration is inadmissible, without asserting anything about whether the underlying claims in a and b happen to be true.*

8.4 Bind

Bind(a, b) records a coupling. Chapters 4 and 5 already gave this operation its most fully worked-out algebraic content: binding x and y under a quadratic evaluation Q produces a bilinear correction $\Delta_Q(x, y) = 2B(x, y)$, exactly recoverable via polarization. The vocabulary of worlds generalizes this: Bind(a, b) need not be numerical at all, and the relational contribution it introduces need not be a bilinear form, or even a scalar. What is preserved from the algebraic case is the two-part structure: a Bind operation is a semantic act, recorded in H, that a and b are now coupled, and a chosen evaluation (a Q, a B, or some other σ) is what measures, after the fact, what that coupling amounts to. Chapter 4's Proposition 4.4 is the special case of this general pattern in which the coupling has already been fixed as bilinear and the measurement is the correction formula itself. Later chapters, particularly Chapter 10, examine what can be inferred about a Bind operation from a measured correction, and what cannot.

8.5 Collapse

Collapse$_\sigma(H)$ is the operation this book has been performing, informally, since the moment the trinomial square was first written as a single number rather than as nine ordered products. It is the only one of the four operations to which Proposition 7.4 directly applies.

Definition 8.4. *A collapse* Collapse$_\sigma$ *is* admissible relative to a task $\mathcal{T}$ *if it does not discard any option-space distinction that $\mathcal{T}$ requires. Making this precise requires a way to say which distinctions in Ω a given task actually depends on, since most tasks care about some property or coarsening of the option space rather than the option space in its full literal detail.*

Definition 8.5. *A* task-observation map *for a task $\mathcal{T}$ is a map $r_\mathcal{T} : \mathcal{O} \to \mathcal{O}_\mathcal{T}$ that retains exactly the distinctions among option spaces relevant to $\mathcal{T}$, and identifies any two option spaces that $\mathcal{T}$ does not need to tell apart. Define*

$$\omega_\mathcal{T} = r_\mathcal{T} \circ \omega : \mathcal{H} \to \mathcal{O}_\mathcal{T}.$$

Corollary 8.6 (Task-relative option-space sufficiency)**.** *Let $\sigma : \mathcal{H} \to \mathcal{S}$ be a derived-state map, let $\omega : \mathcal{H} \to \mathcal{O}$ assign option spaces to histories, and let $r_{\mathcal{T}} : \mathcal{O} \to \mathcal{O}_{\mathcal{T}}$ be a task-observation map for $\mathcal{T}$. Then σ preserves all option-space information relevant to $\mathcal{T}$ if and only if $\omega_{\mathcal{T}} = r_{\mathcal{T}} \circ \omega$ is constant on every fiber of σ. Equivalently, there exists a map $\bar{\omega}_{\mathcal{T}} : \mathcal{S} \to \mathcal{O}_{\mathcal{T}}$ such that*

$$\omega_{\mathcal{T}} = \bar{\omega}_{\mathcal{T}} \circ \sigma.$$

Proof. Apply Proposition 7.4 to the map $\omega_{\mathcal{T}} = r_{\mathcal{T}} \circ \omega$ in place of ω. The required factorization exists exactly when $\omega_{\mathcal{T}}$ is constant on each fiber of σ, which is precisely Proposition 7.4's condition applied to $\omega_{\mathcal{T}}$ rather than to ω itself; nothing in that proposition's proof used any property of ω beyond its being a function out of $\mathcal{H}$, so the substitution is immediate. □

Collapse via σ is therefore admissible relative to $\mathcal{T}$ exactly when

$$\sigma(H_1) = \sigma(H_2) \implies r_{\mathcal{T}}(\omega(H_1)) = r_{\mathcal{T}}(\omega(H_2))$$

for all $H_1, H_2 \in \mathcal{H}$, and its failure has an equally exact witness: a pair of histories with

$$\sigma(H_1) = \sigma(H_2) \quad \text{but} \quad r_{\mathcal{T}}(\omega(H_1)) \neq r_{\mathcal{T}}(\omega(H_2)).$$

This is a deliberate improvement over saying that two histories "agree on every continuation relevant to $\mathcal{T}$," a phrase that leaves relevance to be judged in prose. Relevance is now represented by an explicit map $r_{\mathcal{T}}$, chosen and stated as part of specifying the task, and admissibility becomes a checkable property relative to that stated choice rather than a matter of interpretation. Two people who disagree about whether a given collapse is admissible for a given task are now disagreeing about something locatable: either about what $r_{\mathcal{T}}$ should be, or about whether the stated $r_{\mathcal{T}}$ and σ in fact satisfy the displayed implication. Neither disagreement can be settled by rhetoric alone, which is the entire point of writing the criterion this way.

8.6 Refuse as a constraint on Collapse

Refuse and Collapse are connected in a way worth stating explicitly, since it is easy to treat the four operations as a simple list rather than as a system with real dependencies among them. Refuse does not, by itself, change what any σ computes; a refused pair (a, b) may still have $a \sim_\sigma b$ under some σ's induced relation, if that σ happens not to distinguish them. What Refuse changes is which σ's are subsequently *available* to be applied at all: a Collapse using a σ with $a \sim_\sigma b$ for some refused pair $(a, b) \in R_H$ is not merely inadmissible relative to some task, it is inadmissible outright, by the world's own recorded history, independent of what any downstream task happens to need. This is a stronger and less contingent form of inadmissibility than the task-relative kind defined above, and later chapters (beginning with Chapter 13, on due process as constrained binding) rely on the distinction between the two: some collapses are inadmissible because a particular task needs information they discard, and some are inadmissible full stop, because the world's history has explicitly forbidden the identification they would perform.

Exercises

1. State, in the vocabulary of this chapter, what Pop, Bind, and Collapse operations were performed, in what order, to arrive at the trinomial-square identity of the Preface. Is there a point at which Refuse could sensibly have been invoked but was not?

2. Using Corollary 8.6, write down an explicit $r_{\mathcal{T}}$ for the task "report the total value $2xy$ only," confirm the collection step $xy + yx \mapsto 2xy$ is admissible relative to it, and then write down a different $r_{\mathcal{T}}$, for a task that must distinguish "one product counted twice" from "two products each counted once," relative to which the same collapse is inadmissible.

3. Construct a small example (four or five entries) in which a Refuse(a, b) operation is recorded, and then show a specific σ that would be inadmissible outright as a result, independent of any task.

4. Explain why Bind(a, b), unlike Collapse$_\sigma$, cannot be premature in the sense of Proposition 7.4. What kind of failure, if any, can a Bind operation still be subject to?

Chapter 9

Normal Forms and Noncommutativity

9.1 A default ordering, and why it might not be free

The four operations of Chapter 8, Pop, Refuse, Bind, Collapse, suggest a natural order in which a world is typically built: distinctions are popped, non-equivalences among them are refused, couplings among them are bound, and only then, once the relevant structure has been assembled, is anything collapsed to a derived state. Call this the operational normal form,

$$\text{Pop}^* \to \text{Refuse}^* \to \text{Bind}^* \to \text{Collapse}^*,$$

where the asterisk indicates that any number of applications of that operation may occur in that phase. Every construction in this book so far has, in fact, followed this order: the trinomial square popped x, y, z, bound them pairwise, and only then collapsed the result to $(x + y + z)^2$.

This chapter asks whether that ordering is a genuine constraint, something a world-builder must respect, or merely a convenient habit that happens to match the examples chosen so far. The answer is that it is a real constraint in one direction and not in the other, and the asymmetry is worth stating precisely, because Chapter 10 depends on knowing exactly which reorderings are safe.

9.2 Well-formedness versus commutativity

Some constraints on ordering are definitional rather than substantive. $\text{Bind}(a, b)$ presupposes that a and b have already been popped; there is no meaningful sense in which two entries can be coupled before either is addressable. Call an operation sequence *well-formed* if every Bind is preceded by Pops of both its arguments, and every Refuse is preceded by Pops of both of its arguments. Well-formedness is a syntactic requirement, checkable by inspecting the sequence alone, and it rules out, for instance, $\text{Bind}(a, b)$ appearing before $\text{Pop}(b)$.

Well-formedness is a considerably weaker requirement than the full normal form above. It permits Pop, Refuse, and Bind to be interleaved freely, popping a third entry c after a and b have already been bound, or refusing a pair after some unrelated pair has already been bound, provided each operation's own arguments have been popped first. The question this chapter actually cares about is not whether such interleavings are well-formed, they generally are, but

whether they are *equivalent*: whether performing Pop, Refuse, and Bind operations in a different order, or with a Collapse inserted earlier than the normal form suggests, produces the same eventual state as performing them in the canonical order.

Proposition 9.1 (Factorization through extensional structure)**.** *Let H be a well-formed sequence of Pop, Refuse, and Bind operations, with no Collapse yet applied, and suppose every operation's payload is fixed independently of the order in which it is executed: a Pop's identity, a Refuse's pair, and a Bind's coupling value do not depend on what has already occurred when the operation is performed. Define the* extensional structure *of H as*

$$\tau(H) = \big(P(H),\, R(H),\, B(H)\big),$$

where $P(H)$ is the final set of popped entries, $R(H)$ is the final set of refused pairs, and $B(H)$ is the final set of bound pairs together with their coupling values. If σ factors through τ, that is, $\sigma = \bar{\sigma} \circ \tau$ for some $\bar{\sigma}$, then any well-formed reordering H' of H that preserves the relative order of each Bind or Refuse after the Pops of its own arguments satisfies $\tau(H) = \tau(H')$, and consequently $\sigma(H) = \sigma(H')$.

Proof. Since H and H' are well-formed reorderings of the same underlying operations with order-independent payloads, differing only in sequence, they produce the same final popped set $P(H) = P(H')$ (Pop is a set-insertion operation, and insertion order does not affect the resulting set), the same final refused-pair set $R(H) = R(H')$ (Refuse likewise), and the same final bound-pair set with the same coupling values $B(H) = B(H')$ (Bind likewise, since by hypothesis a coupling value assigned to a pair does not depend on when the pair was bound, only on what value was assigned at the time). Hence $\tau(H) = \tau(H')$, and since $\sigma = \bar{\sigma} \circ \tau$, $\sigma(H) = \bar{\sigma}(\tau(H)) = \bar{\sigma}(\tau(H')) = \sigma(H')$. □

This proposition is doing exactly the same kind of work that Chapter 1 did for ordinary multiplication: it identifies the specific structural reason two different-looking histories produce the same value, namely that the value depends only on the extensional structure $\tau(H)$, a set-like summary of the history, rather than on the sequence itself. Stating it as a factorization through τ, rather than as an informal appeal to "depends only on the final sets," makes visible exactly what is being assumed: that every operation's payload is order-independent (a Bind's coupling value is fixed at the moment of binding and does not later depend on what preceded it) and that σ itself has no sensitivity to anything outside $\tau(H)$. Both halves of this hypothesis are substantive restrictions, and, as the rest of this section shows, many state maps of genuine interest violate one or the other.

9.3 Where the commutativity breaks

Two kinds of σ, and one kind of payload, routinely fail the hypotheses of Proposition 9.1, and it is worth having a concrete instance of each before Collapse is allowed back into the sequence.

The first kind depends on order because a Bind's own payload is not order-independent: if $\mathrm{Bind}(a, b)$ is understood asymmetrically, a influencing b rather than the two being coupled symmetrically, and if the coupling value assigned depends on which entries were already bound at the time (whether a or b was "already settled"), then $B(H)$ itself is no longer a well-defined function of the underlying operations independent of their order, and the extensional-structure hypothesis fails before σ is even considered. Chapter 27 of the original conception of this material

(directed bindings, folded into the applications material in Part IV of this condensed edition) develops this case in full; it is flagged here only to record that Proposition 9.1 explicitly assumed order-independent payloads, and directional coupling with order-dependent values is precisely where that assumption fails.

The second kind depends on order because σ is not merely a function of the final bound-pair set, but of the sequence of intermediate states the world passed through, an evaluation with memory rather than a function of the endpoint alone. A σ that computes, for instance, "the coupling value at the moment the third entry was popped" is sensitive to exactly when that Pop occurred relative to the Binds around it, and reordering the sequence changes its value even though the final sets are identical. Such history-sensitive evaluations are not exotic; Chapter 14 (path dependence and hysteresis, in Part IV) is devoted to them, and the present chapter's role is only to note, in advance, that Proposition 9.1 does not apply to them, precisely because its proof relied on σ factoring through a final, order-blind summary τ.

9.4 Collapse need not commute with subsequent operations

The two failure modes above concern reordering Pop, Refuse, and Bind among themselves, with Collapse held fixed at the end. A sharper question is what happens when Collapse is moved earlier, interleaved with further Pop, Refuse, or Bind operations rather than left until last. The claim to be established is existential, not universal: it is not that no Collapse ever commutes with any later operation (some do, whenever enough structure survives the collapse), but that no general commutation law holds, and a single sound counterexample is enough to establish that.

Proposition 9.2. *There exist histories H_1, H_2, a state map σ, and a subsequent* Bind *operation β such that*

$$\sigma(H_1) = \sigma(H_2) \qquad \textit{but} \qquad \sigma\big(\mathrm{Bind}_\beta(H_1)\big) \neq \sigma\big(\mathrm{Bind}_\beta(H_2)\big).$$

Consequently, the outcome of applying β and then σ is not, in general, determined by $\sigma(H)$ alone: knowing the collapsed state is not sufficient to predict the result of a subsequent binding, even when the subsequent operation itself is fixed and identical in both cases.

Proof. Let H_1 pop two entries with $(x, y) = (1, 2)$, and let H_2 pop two entries with $(x, y) = (-1, -2)$. Let $\sigma(H) = (x + y)^2$, the partial trinomial evaluation after x and y have been bound but before a third entry is introduced. Then

$$\sigma(H_1) = (1 + 2)^2 = 9, \qquad \sigma(H_2) = (-1 - 2)^2 = 9,$$

so $\sigma(H_1) = \sigma(H_2)$. Now let β pop a further entry $z = 1$ and bind it to the existing pair, extending the evaluation to the full trinomial square, $\sigma(\mathrm{Bind}_\beta(H)) = (x + y + z)^2$. Applying this identical β to each history,

$$\sigma\big(\mathrm{Bind}_\beta(H_1)\big) = (1 + 2 + 1)^2 = 16, \qquad \sigma\big(\mathrm{Bind}_\beta(H_2)\big) = (-1 - 2 + 1)^2 = 4,$$

and $16 \neq 4$ despite $\sigma(H_1) = \sigma(H_2)$. □

This is exactly an instance of task-relative option-space failure in the sense of Corollary 8.6, with the task $\mathcal{T}$ being "determine the result of extending by $z = 1$ under the trinomial evaluation." The two histories H_1 and H_2 lie in the same fiber of σ but require different values once β is

applied, which is precisely the witness of inadmissibility that corollary describes: $\sigma(H_1) = \sigma(H_2)$ while $r_{\mathcal{T}}(\omega(H_1)) \neq r_{\mathcal{T}}(\omega(H_2))$, since the two histories admit the same subsequent Bind but yield different results from it. The scalar $\sigma(H) = 9$ does not fail to determine x and y separately, that much was already clear from Chapter 3; what this example adds is that it does not even determine the *sum* $x + y$ up to sign, since $(x + y)^2 = 9$ is satisfied by $x + y = 3$ or $x + y = -3$, and the two cases diverge under a later binding.

This is the formal statement of something Chapter 6 and Chapter 8 both anticipated informally: collapsing early can foreclose determinacy that collapsing late would have kept available, and Proposition 9.2 shows that this is a genuine possibility for Collapse in general, not a special pathology confined to one contrived example, even though establishing it rigorously required exhibiting one particular case rather than proving a fully universal failure. The operational normal form, Pop and Refuse and Bind before Collapse, is therefore not merely a convention this book has followed for expository convenience. Following it is one reliable way to keep every later Bind's outcome determined by the retained history; departing from it, collapsing before a later Bind is applied, does not guarantee failure, since Proposition 9.1's hypothesis can still be met by a sufficiently well-behaved σ, but it removes the guarantee, and Proposition 9.2 shows the guarantee can genuinely fail rather than merely seeming, on casual inspection, like it might.

9.5 Binding before revision

A closely related asymmetry deserves a name of its own, since it recurs across the applications in Part IV under different vocabulary each time. Let U denote an update or revision operation applied to an individual entry, for instance replacing x with a corrected value x'. The question is whether

$$U\big(\mathrm{Bind}(x, y)\big) \quad \text{and} \quad \mathrm{Bind}\big(U(x), U(y)\big)$$

agree, where the left side updates the entries after they have been bound and the right side updates them separately before binding.

Example 9.3. *Let $Q(x, y) = (x + y)^2$ and let U replace x with $x' = x + \delta$ for some correction δ, leaving y unchanged, so that on the right side, $U(y) = y$. Then*

$$Q\big(U(x), U(y)\big) = (x + \delta + y)^2 = (x + y)^2 + 2\delta(x + y) + \delta^2.$$

If instead the correction is applied to the already-bound quantity $s = x + y$ directly, treating U as "add δ to the bound total," the result is $(s + \delta)^2 = (x + y + \delta)^2$, the same expression in this case, because addition of a scalar correction happens to commute with the specific binding operation (ordinary addition) used here. This agreement is a feature of this example, not a general fact: if U instead corrected only the relationship *between x and y, for instance rescaling the cross-term's coupling from 1 to some $c \neq 1$ without touching x or y individually, no operation on the bound scalar $s = x + y$ alone could reproduce it, since s no longer records x and y separately.*

The general lesson, stated carefully rather than overclaimed from one example: revising components separately and revising a bound whole are the same operation only when the update happens to respect the specific way the components were combined, and there is no general guarantee of this. A representation that has already collapsed x and y into $s = x + y$, a genuine many-to-one projection rather than an invertible relabeling, cannot, in general, be corrected in a way that is equivalent to correcting x and y individually and rebinding them, for a reason of the same shape as Proposition 9.2: the collapsed representation s does not retain the individually addressable entries some corrections would need to act on.

9.6 Summary, and what remains

This chapter has drawn three related but distinct conclusions. Pop, Refuse, and Bind commute freely among themselves whenever every operation's payload is order-independent and the eventual state map factors through the resulting extensional structure $\tau(H)$ (Proposition 9.1), and fail to commute exactly when a payload is itself order-dependent (directional coupling) or when the state map is sensitive to intermediate history rather than only to $\tau(H)$, cases flagged here for treatment in Part IV rather than resolved. Collapse need not commute with a subsequent Bind: there exist a collapse and a later binding for which the eventual value depends on more than the collapsed state alone, so that two histories collapsing to the same state can diverge once the same further operation is applied to each (Proposition 9.2). And revising a bound whole is not generally equivalent to revising its components and rebinding them, for a closely related reason.

What remains open, and is the subject of Chapter 10, is a question this chapter has assumed rather than answered: given only a measured or observed correction, a Δ_Q or its more general analogue, what can actually be inferred about whether a Bind operation occurred at all, as opposed to the correction arising from some other source. Chapter 5 raised a version of this question for orthogonality and vanishing corrections specifically; Chapter 10 generalizes it to the full operational vocabulary of this chapter.

Exercises

1. Construct a well-formed sequence of two Pops and two Binds, and a reordering of it, such that both produce the same result under a σ that sums all bound-pair coupling values. Then modify σ so that it depends on the order in which the Binds occurred, and show the two histories now disagree.

2. In the directed-binding failure mode described in this chapter, construct a minimal three-entry example where reversing the order of two asymmetric Binds changes the final coupling value, and identify exactly which step in the proof of Proposition 9.1 this example violates.

3. Using the sign-ambiguity example from the proof of Proposition 9.2, determine precisely what additional information, beyond the scalar $\sigma(H) = (x+y)^2$, would need to be retained so that the result of a later Bind of $z = 1$ under the trinomial evaluation becomes computable directly from the retained information rather than requiring the original history H. (Consider whether retaining the signed sum $x + y$, rather than its square, is enough, and whether anything less than that suffices.)

4. Give an example, from outside mathematics, of an update that is applied correctly to a bound or aggregated quantity but would have given a different, and more correct, result if it had instead been applied to the individual components before they were aggregated.

Chapter 10

The Hierarchy of Transformation, and the Cross-Term as Witness

10.1 A hierarchy, made explicit

The corrections to Chapters 6 through 9 were all instances of a single distinction that has not, until now, been given a name of its own: the difference between changing how a history is represented and losing access to what the history contains. It is worth stating the resulting hierarchy plainly before using it, rather than leaving it to be reconstructed piecemeal from four chapters of separate repairs.

$$\text{re-description} \subsetneq \text{transformation} \supseteq \text{projection/identification} \supseteq \text{Collapse relative to a task } q.$$

A *transformation* is simply any map $T : \mathcal{H} \to \mathcal{D}$ out of the space of histories, into some representation domain $\mathcal{D}$, which may itself be another space of histories, a coordinate system, a scalar, or anything else. Nothing has been said yet about whether T preserves anything; this is the widest category, and everything else in the hierarchy is a special case of it.

A transformation is a *re-description* if it is injective and its inverse, or the specific means of computing it (an explicit matrix P, an explicit decoding procedure), is retained alongside $T(H)$. Chapter 6's corrected treatment of diagonalization is the running example: $\mathbf{w} = P^{\mathsf{T}}\mathbf{x}$ together with the retained P is a re-description of $\mathbf{x}$, not a loss of anything, because $\mathbf{x} = P\mathbf{w}$ recovers the original exactly.

A transformation is a *projection or identification* if it is not injective, merging at least two distinct histories, *or* if it is injective but its inverse is not retained or not practically computable, the effective-irrecoverability case flagged already in Chapter 7. Both cases share the property that, in practice, something about H is no longer extractable from $T(H)$ alone, whether because it was genuinely erased by the merge or because the means of un-merging it was discarded separately.

Collapse relative to a task q, finally, is not a further restriction on T itself, but a claim about the pair (T, q): it names those projections or identifications that specifically lose information task q needed. This is why Collapse sits inside projection/identification in the diagram rather than being identified with it: not every projection is a Collapse relative to every task, since a given loss of information may simply be irrelevant to what a particular q asks for. Noninvertibility, by itself, is not synonymous with Collapse. It becomes Collapse only once a task is named

that the noninvertibility happens to defeat, and Section 10.4 below returns to the one important exception, the task that demands recovery of everything, for which noninvertibility and Collapse do coincide.

10.2 The general sufficiency criterion

Corollary 8.6 was stated specifically for a state map $\sigma : \mathcal{H} \to \mathcal{S}$, in the setting where Chapter 8 had already fixed σ as the operator performing Collapse. The definition of transformation just given is broader, allowing T to map into any domain $\mathcal{D}$, not only a designated state space, and it is worth restating the sufficiency criterion at this level of generality, since the applications later in this chapter need it in exactly this form.

Definition 10.1. *Let $T : \mathcal{H} \to \mathcal{D}$ be a transformation, let $\omega : \mathcal{H} \to \mathcal{O}$ assign option spaces to histories, and let $r_q : \mathcal{O} \to \mathcal{O}_q$ be an observation map for a task q, retaining exactly the distinctions relevant to q. Say T is* q-sufficient *if there exists $\bar{r} : \mathcal{D} \to \mathcal{O}_q$ with*

$$r_q \circ \omega = \bar{r} \circ T.$$

Theorem 10.2 (General sufficiency criterion)**.** *T is q-sufficient if and only if $T(H_1) = T(H_2)$ implies $r_q(\omega(H_1)) = r_q(\omega(H_2))$, for all $H_1, H_2 \in \mathcal{H}$.*

Proof. Identical to the proof of Corollary 8.6, with σ replaced by the general transformation T; nothing in that argument used any property of σ beyond its being a function out of $\mathcal{H}$. □

Proposition 10.3. *Every re-description is q-sufficient for every task q.*

Proof. If T is a re-description, it is injective, so $T(H_1) = T(H_2)$ implies $H_1 = H_2$, which trivially implies $r_q(\omega(H_1)) = r_q(\omega(H_2))$ for any r_q and any q. Theorem 10.2 then gives q-sufficiency. □

Proposition 10.3 is the formal statement of the entire point of the Chapter 6 correction: an invertible change of coordinates, with the inverse retained, can never be Collapse relative to *any* task, however demanding, because it never merges distinct histories in the first place. Every genuine failure of sufficiency, and hence every genuine instance of Collapse, occurs strictly within the projection/identification region of the hierarchy, among transformations that are either non-injective or have had their inverse discarded.

10.3 Collapse is relative, except in one case

It might seem that a sufficiently demanding task would make almost any projection into a Collapse, and this is close to correct, but it is worth stating exactly how close.

Definition 10.4. *Let q^* denote the* maximal task*: the task whose observation map $r_{q^*} : \mathcal{O} \to \mathcal{O}$ is the identity, retaining every distinction ω can register at all.*

Corollary 10.5. *T is q^*-sufficient if and only if $T(H_1) = T(H_2)$ implies $\omega(H_1) = \omega(H_2)$. In particular, if T is injective, T is automatically q^*-sufficient; and if ω itself is injective wherever T is not (that is, if merged histories under T never happen to share an option space), T fails to be q^*-sufficient exactly when T is non-injective.*

The second clause is the qualification worth dwelling on. It is not, in general, true that non-injectivity of T alone guarantees failure of q^*-sufficiency, because two histories can be merged by T and still happen to have identical option spaces, exactly as Chapter 7's option-sufficient-but-non-injective account-balance example showed. What *is* true is the converse relationship this book has been assuming informally: an absolute, task-independent notion of "lossless" collapses onto the special case $q = q^*$, and adopting that absolute notion is a choice, not a default the algebra forces. Most of the applications in Part IV will use a specific, less demanding q, tailored to the task at hand, precisely because most real questions do not require recovering everything, only the specific continuation the task in question depends on.

10.4 The sign-ambiguity example, in the general vocabulary

Proposition 9.2 showed, by direct construction, that $\sigma(H) = (x + y)^2$ fails to determine the result of a subsequent Bind with $z = 1$, using two histories, $(x, y) = (1, 2)$ and $(x, y) = (-1, -2)$, sharing $\sigma(H) = 9$. Restated in this chapter's vocabulary: let $T = \sigma$, let q_1 be the task "report $(x + y)^2$," and let q_2 be the task "report the result of extending by $z = 1$ under the trinomial evaluation." Then T is q_1-sufficient, trivially, since T literally computes what q_1 asks for; and T is not q_2-sufficient, by the explicit counterexample already given. The same transformation, applied to the same pair of histories, is a harmless re-description of the relevant fact for one task and a genuine Collapse for the other. This is the clean illustration Corollary 8.6 promised in the abstract, made concrete: adequacy is a property of (T, q) together, never of T alone, and a single transformation routinely sits on both sides of that relation depending only on which question is subsequently asked of it.

10.5 Is the cross-term a witness of Bind?

Chapter 8 drew a careful distinction, when Bind was first defined, between the semantic act of binding, a specific entry recorded in H, and the numerical evaluation, a Q or a B, that measures what that act amounts to. That distinction can now be posed as a precise sufficiency question, closing a line raised as early as Chapter 5 and left open since.

Let the task q_{bind} be: *determine whether* $\text{Bind}(x, y)$ *was recorded in* H, that is, whether $(x, y) \in B(H)$ in the notation of Chapter 9. Let T_Δ be the transformation that reads off the correction $\Delta_Q(x, y)$ for two popped entries x, y, computed directly from their values, independently of whether a Bind operation was ever recorded between them.

Proposition 10.6. *T_Δ is not q_{bind}-sufficient.*

Proof. It suffices to exhibit two histories agreeing on T_Δ but disagreeing on whether $(x, y) \in B(H)$. Let both H_1 and H_2 pop the same values, $x = 1$, $y = 1$. In H_1, no Bind operation is recorded between them at all. In H_2, $\text{Bind}(x, y)$ is explicitly recorded, with some coupling assigned. Recording a Bind operation is an addition to the operational history; it does not alter the popped values of x and y themselves. Since T_Δ is defined as a function of popped values alone, computing $\Delta_Q(x, y) = 2xy$ directly from $x = 1, y = 1$ regardless of what has or has not been recorded in $B(H)$, $T_\Delta(H_1) = T_\Delta(H_2) = 2 \cdot 1 \cdot 1 = 2$. This value is nonzero, so the two histories agree on a genuine, nontrivial correction, not merely on an accidental zero; yet $(x, y) \notin B(H_1)$ while $(x, y) \in B(H_2)$. □

The construction in the proof is almost too simple to feel like it settles anything, and that is itself informative. T_Δ is defined as a function of x and y's *values*, and Chapter 8 was explicit that a Bind operation is a fact about H, not a fact that changes what x and y evaluate to. A correction computed purely from values can therefore never, by construction, distinguish a history in which a Bind was recorded from one in which it was not, whenever the two histories happen to share the same popped values, regardless of what those values are or what Q is chosen. This is a sharper and more general statement than anything Chapter 5's four-fold list could offer, because it does not depend on Δ_Q taking any particular value, such as zero; the proof used $\Delta_Q(1,1) = 2$, a genuinely nonzero correction, precisely to make this point, and the same argument goes through at any other pair of values, since recording a Bind operation is exactly the kind of operational fact that, by design, does not itself alter what is being bound.

The natural next question is whether the converse direction fares any better: does $\Delta_Q(x,y) \neq 0$ at least entail that a Bind was recorded? It does not, for the same reason in reverse. Nothing in the definition of Q requires x and y to have been the subject of any Bind operation before $Q(x+y)$ can be evaluated; Q is defined for any two popped, addressable entries, whether or not the world's history ever explicitly coupled them. Two entries that were merely popped into the same history, with no Bind ever recorded, can be evaluated by an external Q that happens to produce a nonzero correction anyway. The correction Δ_Q, computed this way, answers a question about the evaluation Q, not a question about the operational history H, and Proposition 10.6 together with this observation shows that the two questions are simply not the same question, in either direction.

10.6 What would make it a witness

This need not be the end of the matter, and it is worth being constructive about what would repair it, since the applications in Part IV frequently do want a correction to serve as evidence of an operational fact. Recall Chapter 9's extensional structure, $\tau(H) = (P(H), R(H), B(H))$, where $B(H)$ is specifically the set of *recorded* bound pairs with their coupling values, as opposed to T_Δ, which reads off a correction from popped values directly and independently of $B(H)$.

Proposition 10.7. *Let σ_B be a transformation that factors through τ and depends on a pair (x,y) only via its entry in $B(H)$, assigning zero contribution to any pair not in $B(H)$, regardless of that pair's popped values. Then observing a nonzero contribution attributable to (x,y) under σ_B implies $(x,y) \in B(H)$, that is, implies a Bind was recorded between x and y.*

Proof. Immediate from the hypothesis: σ_B is constructed so that its value depends on (x,y) only through membership in $B(H)$, assigning zero whenever $(x,y) \notin B(H)$. A nonzero attributed contribution is therefore only possible when $(x,y) \in B(H)$. □

The difference between T_Δ, which fails Proposition 10.6, and σ_B, which succeeds by Proposition 10.7, is not a difference in mathematical sophistication. It is a difference in what each transformation is defined to consult: T_Δ consults only the popped values, blind to whether a Bind was ever recorded, while σ_B consults the operational record $B(H)$ directly, by construction. Whether a measured correction is legitimately a witness of binding, in any of the applications from Chapter 12 onward, therefore turns entirely on which of these two kinds of evaluation is actually in use, a distinction easy to blur in prose ("the correlation shows they interacted") and now, at least, precise to check: does the evaluation read the operational record, or does it

merely compute a function of values that happen to coexist in the same history regardless of what operations bound them.

10.7 Closing Part III

Part III set out to make explicit the operations this book had been performing informally since the Preface. Chapter 7 supplied the vocabulary of worlds, histories, and option spaces, and the precise sense in which a derived state can lose exactly the continuations a task still needs (Proposition 7.4, Corollary 8.6). Chapter 8 named four operations, three purely constructive and one, Collapse, capable of discarding what the other three built. Chapter 9 showed when reordering the constructive operations is safe and when Collapse fails to commute with what comes after it. This chapter has generalized that machinery beyond Collapse specifically, to any transformation of a history whatsoever, organized the results into an explicit hierarchy running from lossless re-description to task-relative Collapse, and used the resulting theorem to settle, in the negative, whether an algebraic correction can by itself serve as evidence of an operational binding, while also showing exactly what a correction would need to consult in order to serve that role legitimately.

Part IV turns from this general apparatus to specific substrates: graphs, probability, dynamics, computation, evidence, and value, each of which will turn out to instantiate the same hierarchy, the same sufficiency question, and, in more than one case, the same trap of mistaking a value-level correction for a witness of an operational fact it was never built to detect.

Exercises

1. Confirm directly that the account-balance transformation from Chapter 7 is a projection/identification (non-injective) that is nonetheless q-sufficient for the task "report current funds," and identify a second task, distinct from the three-withdrawals rule already given, for which it fails to be sufficient.

2. Construct an example of a transformation that is injective (hence, in principle, a candidate for re-description) but for which the inverse is not retained, and confirm it is q^*-sufficient in principle (Corollary 10.5) while failing to be practically q^*-sufficient under a stated resource bound. Relate this to Chapter 7's distinction between informational and effective recoverability.

3. Extend the proof of Proposition 10.6 to show that no transformation defined purely as a function of popped values, without consulting $B(H)$, can ever be q_{bind}-sufficient, regardless of which Q is chosen.

4. Construct a σ_B satisfying the hypothesis of Proposition 10.7 explicitly, for a world with three popped entries and two recorded Binds, and confirm by direct computation that it correctly reports which pairs were bound and assigns zero to the unbound pair regardless of that pair's popped values.

Part III

Graphs, Probability, and Information

Chapter 11

Graphs and Binding Matrices

11.1 A standing convention

Before turning to specific substrates, one distinction from Chapter 10 needs to be adopted as a standing convention for everything that follows, precisely because it is the distinction most likely to be blurred by casual language once concrete applications are in view. *Many-to-one* is a structural property of a transformation T: it says only that T is not injective, that at least two distinct histories share an image under T. *Collapse* is a relational property, holding or failing of a pair (T, q): it says that T merges histories which some named task q still needs to tell apart. These are not the same claim, and the gap between them is not a technicality.

Remark 11.1 (Non-injectivity supplies possibility, not automatic failure)**.** *A non-injective T is a* candidate *for Collapse relative to some task, since it merges histories that some sufficiently demanding q might need distinguished. It is not automatically a Collapse relative to* every *task. By Theorem 10.2, T fails q-sufficiency only when a merged fiber of T contains histories H_1, H_2 with $r_q(\omega(H_1)) \neq r_q(\omega(H_2))$; if ω (or its r_q-observed image) happens to be constant across every fiber T merges, T remains fully q-sufficient despite being many-to-one, exactly as Chapter 7's account-balance map was option-sufficient for "report current funds" despite merging infinitely many deposit histories.*

The temptation this remark is meant to guard against is specific to the chapters ahead. A graph quotient, a summary statistic, or a communication channel all merge distinct inputs into a common output by their very definition, and it will be easy to describe any of them as "a collapse" on that basis alone. The correct statement is narrower: each is a many-to-one transformation, a candidate for Collapse, and whether it actually is one depends on which task is subsequently asked of it. This chapter, and the two after it, will repeatedly name the task explicitly rather than leaving "collapse" to do unexamined work.

11.2 A translation across three substrates

The general sufficiency criterion of Chapter 10 has the same shape regardless of what $\mathcal{H}$, $\mathcal{D}$, and q turn out to mean in a given substrate, and it is worth naming the translation once, here, rather than re-deriving it three times.

In a *graph*, the history is the graph itself, with its vertices and weighted edges; a transformation is an operation such as relabeling, quotienting, aggregating, or contracting vertices; and the

question this chapter asks is which structural distinctions, which pairs of vertices, which edge weights, which paths, survive a given such operation, relative to a stated downstream use of the graph.

In *probability*, taken up in Chapter 12, the history is the full joint distribution or the underlying random mechanism; a transformation is a statistic, a function of the data; and the question is which statistics are *sufficient* for a target variable or a decision, in exactly the technical sense that phrase already carries in statistics, and in exactly the sense Theorem 10.2 generalizes.

In *information theory*, taken up in Chapter 13, the history is the source signal; a transformation is a channel or an encoding; and the question is which channels preserve the distinctions a decoder, or a task downstream of the decoder, actually needs.

These are not three independent analogies to the algebra of Parts I through III. They are three instantiations of the same theorem, with $\mathcal{H}$, T, and q given substrate-specific names, and the point of saying so before working through any of them is that a genuine result, obtained honestly for graphs, will transfer to probability and information without needing to be re-argued from scratch, once the correspondence between the vocabularies has been made explicit rather than left implicit.

11.3 Vertices, edges, and the binding matrix

A graph on n vertices, with a symmetric weight matrix $A = (a_{ij})$ recording the strength of the edge between vertices i and j (and a_{ii} recording a self-weight or self-loop, where relevant), is formally identical to the quadratic form apparatus of Chapter 4. Nothing new needs to be built: the vertices are the popped entries $x_1, \ldots, x_n$, the edges are the pairwise couplings, and

$$Q(\mathbf{x}) = \mathbf{x}^{\mathsf{T}} A \mathbf{x} = \sum_i a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j$$

is one natural evaluation of the graph, with a_{ij} recoverable from Q alone via polarization exactly as in Chapter 5. Calling A a *binding matrix* rather than merely an adjacency or weight matrix is a choice of emphasis, not a new object: it is the same matrix, read as a record of pairwise binding corrections rather than as a table of distances or capacities.

11.4 Directed bindings

Chapter 9 flagged, and deliberately deferred, the case where a Bind operation is asymmetric: $\mathrm{Bind}(a, b)$ influencing a from b in a way not symmetric in the two arguments. Graphs are the natural setting to redeem that promise. A *directed* graph has a weight matrix A that need not be symmetric, a_{ij} representing the edge from i to j and a_{ji}, possibly different, the edge from j to i. The bilinear form $B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathsf{T}} A \mathbf{y}$ still makes sense for non-symmetric A, but Proposition 4.4 no longer holds in its stated form, since its proof used $B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x})$ explicitly. What survives is a weaker identity: writing $A = A_S + A_K$ for the symmetric part $A_S = \frac{1}{2}(A + A^{\mathsf{T}})$ and antisymmetric part $A_K = \frac{1}{2}(A - A^{\mathsf{T}})$,

$$\mathbf{x}^{\mathsf{T}} A \mathbf{x} = \mathbf{x}^{\mathsf{T}} A_S \mathbf{x},$$

since $\mathbf{x}^{\mathsf{T}} A_K \mathbf{x} = 0$ for any antisymmetric A_K (a direct computation: $\mathbf{x}^{\mathsf{T}} A_K \mathbf{x}$ equals its own negative, transposing a scalar). This is a precise, and slightly unsettling, fact: the quadratic

form $Q(\mathbf{x}) = \mathbf{x}^\mathsf{T} A\mathbf{x}$ is entirely blind to the antisymmetric part of A, no matter how large that part is. A directed graph's asymmetry, the entire content of "which direction the edge points," is invisible to any evaluation built as a single quadratic form of the vertex values. Recovering directional information requires either a genuinely bilinear evaluation $B(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} \neq \mathbf{y}$ representing two different roles (a source vector and a target vector, rather than one vector paired with itself), or an evaluation that reads the matrix A directly rather than only Q's scalar output. This is the graph-theoretic sharpening of Chapter 10's closing distinction between T_Δ and σ_B: a scalar quadratic evaluation of a directed graph is, by construction, not q-sufficient for any task q that asks about direction, however cleverly it is defined, because the antisymmetric part is annihilated identically, not merely obscured.

11.5 Sparse binding and orthogonality

Most graphs of practical interest have $a_{ij} = 0$ for most pairs (i, j): a sparse binding matrix, most pairs of vertices simply not connected. This is not a new phenomenon so much as the graph-theoretic name for something Chapter 5 already characterized precisely: $a_{ij} = 0$ is exactly $B(\mathbf{e}_i, \mathbf{e}_j) = 0$, orthogonality of the i-th and j-th coordinate directions with respect to Q. Every qualification Chapter 5 attached to a vanishing correction applies here without modification. An absent edge between i and j in a graph built from a specific evaluation Q does not, by itself, certify that i and j are unrelated in whatever the vertices represent; it certifies only that this particular Q registers no coupling between them, and Chapter 5's four-fold list of alternative explanations, genuine independence, an undetectable form of dependence, cancellation of several couplings, or a Q too coarse to register the relation, transfers directly. A common practical error is to read a sparse graph, constructed from one specific measured or modeled interaction, as evidence of the absence of any relationship at all; the algebra has never licensed that inference, in graphs or anywhere else.

11.6 Quotients as transformations

The concept this chapter has been building toward is the graph quotient: an operation that merges several vertices into one, replacing x_i and x_j (and possibly others) with a single new vertex $\bar{x}$, and redistributing or aggregating the edges accordingly. A quotient is a transformation $T : \mathcal{H} \to \mathcal{D}$ in exactly Chapter 10's sense, mapping a graph to a smaller graph, and Remark 11.1 applies immediately: a quotient is many-to-one by construction (every quotient that actually merges vertices is non-injective on graphs distinguished only by what happens inside the merged group), but whether it constitutes Collapse depends entirely on the task.

Example 11.2. *Let a graph have four vertices,* $\{1, 2, 3, 4\}$*, and let a quotient* T *merge* 1 *and* 2 *into a single vertex* $\bar{12}$*, summing their edges to* 3 *and to* 4 *and discarding whatever edge weight existed between* 1 *and* 2 *themselves. For the task* q_1*, "report the total edge weight leaving the group* $\{1, 2\}$ *toward vertex* 3*," T is* q_1*-sufficient: the quotient is built to compute exactly this sum, and any two original graphs agreeing on that sum will agree after quotienting, by construction. For the task* q_2*, "report the edge weight between* 1 *and* 2 *specifically," T is not* q_2*-sufficient: that weight is exactly what the quotient discarded, and two graphs differing only in the* 1–2 *edge weight, with every other weight identical, map to the identical quotient graph.*

This is the graph-theoretic instance of Corollary 8.6, and it is worth stating the general principle the example illustrates, since it recurs in every quotient construction to follow. A quotient is defined by which edges it aggregates and which it discards; it is q-sufficient exactly for those tasks whose answer depends only on the aggregated quantities the quotient retains, and it fails exactly for those tasks that need something the aggregation summed away or dropped. Designing a quotient, in this sense, is not a matter of finding "the" correct coarsening of a graph; it is a matter of naming the task first and building the coarsening to preserve what that task needs, which is Chapter 8's admissibility criterion for Collapse, applied here to a specific and highly visual kind of transformation.

11.7 Toward hyperedges

One further loose end from Part I is worth naming here, even though its full treatment lies beyond this condensed edition's scope. Chapter 2's closing exercise noted, without resolving, that the additivity of pairwise corrections holding for $Q(x) = x^2$ fails for $Q(x) = x^3$, whose expansion contains a genuine three-way term, $6xyz$, not decomposable into any sum of pairwise contributions. In graph language, a binding matrix A captures only pairwise structure; a system whose evaluation contains an irreducible three-way term requires a *hyperedge*, a single relation spanning three or more vertices at once, not representable as a sum of ordinary edges no matter how the pairwise weights are chosen. Recognizing when a system needs a hyperedge, rather than merely a denser ordinary graph, is exactly the question of whether its governing evaluation Q decomposes into self-terms and pairwise terms alone, as the trinomial square does, or requires terms of higher arity, as the cubic does. This condensed edition does not develop hypergraph theory at length, but the reader should leave this chapter able to recognize the diagnostic: a pairwise binding matrix, however finely weighted, is the wrong tool exactly when the underlying evaluation's binding correction fails to decompose pairwise, and Chapter 2's cubic exercise was the first, and remains the simplest, witness of that failure.

Exercises

1. Confirm by direct computation that $\mathbf{x}^{\mathsf{T}} A_K \mathbf{x} = 0$ for the antisymmetric matrix $A_K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and any $\mathbf{x} = (x, y)$, and relate this to the general argument given in the text.

2. Construct a 3×3 sparse binding matrix (at least two off-diagonal entries zero) and identify a task for which the resulting graph is q-sufficient despite the sparsity, and a task for which it is not.

3. Extend the quotient example of this chapter to a five-vertex graph, merging three vertices into one, and state explicitly which of the ten possible pairwise edge weights in the original graph are retained, aggregated, or discarded by the quotient.

4. Using Remark 11.1, construct a non-injective graph quotient that is nonetheless q-sufficient for a nontrivial task of your choosing, in the style of Chapter 7's account-balance example.

Chapter 12

Variance, Sufficiency, and What a Statistic Retains

12.1 Variance as the cleanest instance

Of every substrate this book will visit, none instantiates the quadratic-form apparatus of Part II more directly than ordinary variance. Let $X_1, \dots, X_n$ be random variables, and consider their sum $S = \sum_i X_i$. Variance is not merely analogous to a quadratic form; it literally is one, applied to the vector of coefficients in a linear combination, and the identity this section derives is the binding correction of Chapter 2, unchanged, appearing under a different name.

Proposition 12.1. *For random variables $X_1, \dots, X_n$ with finite variance,*

$$\operatorname{Var}\left(\sum_i X_i\right) = \sum_i \operatorname{Var}(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j).$$

Proof. Write $\mu_i = \mathbb{E}[X_i]$. Then

$$\operatorname{Var}\left(\sum_i X_i\right) = \mathbb{E}\left[\left(\sum_i (X_i - \mu_i)\right)^2\right] = \mathbb{E}\left[\sum_i (X_i - \mu_i)^2 + 2\sum_{i<j} (X_i - \mu_i)(X_j - \mu_j)\right],$$

using exactly the trinomial-square expansion of Chapter 1, applied to the centered variables $X_i - \mu_i$ in place of x_i. Taking expectations term by term gives $\sum_i \operatorname{Var}(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j)$, since $\mathbb{E}[(X_i - \mu_i)^2] = \operatorname{Var}(X_i)$ and $\mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = \operatorname{Cov}(X_i, X_j)$ by definition. □

The identification with Chapter 4's apparatus is exact, not approximate. Let $Q(\mathbf{x}) = \operatorname{Var}(\sum_i x_i X_i)$ for coefficients $\mathbf{x} = (x_1, \dots, x_n)$; this is a quadratic form in $\mathbf{x}$ with matrix $A_{ij} = \operatorname{Cov}(X_i, X_j)$ (so $A_{ii} = \operatorname{Var}(X_i)$), and Proposition 12.1 is simply this Q evaluated at $\mathbf{x} = (1, \dots, 1)$. Polarization (Chapter 5) recovers $\operatorname{Cov}(X_i, X_j)$ from evaluations of Q alone, exactly as it recovered any other bilinear coupling, and diagonalizing A (Chapter 6) produces uncorrelated linear combinations of the X_i, the statistical name for an eigenbasis in which the covariance structure has been diagonalized rather than removed, with every qualification Chapter 6 attached to that operation applying without change: the uncorrelated combinations are a re-description of the same random vector, not a description of a system that was never correlated.

12.2 Independence, uncorrelatedness, and hidden dependence

Chapter 5 warned, in the abstract, that a vanishing bilinear correction is compatible with several distinct underlying situations, only one of which is genuine independence. Probability is where that warning has its sharpest and most standard illustration, worth working through concretely rather than left as a general caution.

Example 12.2. *Let X be uniformly distributed on $[-1, 1]$, and let $Y = X^2$. Then*

$$\mathbb{E}[X] = 0, \qquad \mathbb{E}[XY] = \mathbb{E}[X^3] = \int_{-1}^{1} \tfrac{1}{2}x^3 \, dx = 0,$$

so $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 - 0 = 0$. Yet Y is a deterministic function of X: knowing X determines Y exactly, the strongest possible form of dependence, not the weakest.

This is Chapter 5's second explanation, "a dependence this specific Q is not built to detect," realized exactly: the quadratic evaluation $Q(\mathbf{x}) = \mathrm{Var}(\sum x_i X_i)$, and the covariance it recovers via polarization, is sensitive only to *linear* association. A relationship that is perfectly strong but symmetric around a vanishing linear trend, as $Y = X^2$ is symmetric around $X = 0$, produces a bilinear correction of exactly zero, for the same structural reason that Chapter 5's orthogonality was shown to be a fact about B specifically, not a fact about the underlying objects in general. The lesson transfers without modification: $\mathrm{Cov}(X, Y) = 0$ certifies the absence of linear association and nothing more; treating it as certifying general independence is exactly the error Chapter 5 named and Example 12.2 makes concrete.

12.3 Portfolio risk

A weighted version of Proposition 12.1 is the standard setting in which this identity is used to make decisions rather than merely to compute a number. Let $w_1, \ldots, w_n$ be portfolio weights (fractions of a total investment) allocated to assets with returns $X_1, \ldots, X_n$. The portfolio return is $\sum_i w_i X_i$, and its variance, the standard measure of portfolio risk, is

$$\sigma_P^2 = \mathrm{Var}\left(\sum_i w_i X_i\right) = \sum_i w_i^2 \,\mathrm{Var}(X_i) + 2\sum_{i<j} w_i w_j \,\mathrm{Cov}(X_i, X_j),$$

an immediate instance of $Q(\mathbf{w}) = \mathbf{w}^{\mathsf{T}} A \mathbf{w}$ for $A_{ij} = \mathrm{Cov}(X_i, X_j)$, exactly as before, now evaluated at the weight vector $\mathbf{w}$ rather than at $(1, \ldots, 1)$. The entire content of diversification, the empirical fact that a portfolio can have lower risk than any of its individual components, is a direct consequence of the cross-terms in this expansion: if $\mathrm{Cov}(X_i, X_j) < 0$ for some pair, the corresponding term in σ_P^2 is negative, and the portfolio's variance can fall below $\sum_i w_i^2 \,\mathrm{Var}(X_i)$, the value the separate evaluation of Chapter 2 would have predicted. This is the binding correction Δ_Q, from Chapter 2 onward, appearing with a genuinely negative sign in a case that matters economically, not merely as a possibility flagged in the abstract discussion of Chapter 2's Section on sign.

12.4 Sufficient statistics

The word "sufficient" has had a specific technical meaning in statistics for close to a century, independently of anything in this book, and it is worth making the connection to Theorem 10.2

explicit rather than merely gesturing at a resemblance, since the classical notion turns out to be a precise instance of the general one, not merely an analogy to it.

Definition 12.3. *Let $X = (X_1, \dots, X_n)$ be data drawn from a distribution depending on a parameter θ. A statistic $T(X)$ is* sufficient for θ *if the conditional distribution of X given $T(X)$ does not depend on θ.*

Theorem 12.4 (Fisher–Neyman factorization, stated without proof)**.** *$T(X)$ is sufficient for θ if and only if the likelihood factors as*

$$f(x;\theta) = g\big(T(x),\theta\big)\, h(x)$$

for some functions g and h, with h not depending on θ.

To connect this to Theorem 10.2, identify the "history" $\mathcal{H}$ with the space of possible data samples x, and let the task q_θ be: *recover the likelihood function $f(x;\theta)$ as a function of θ, up to a factor not depending on θ.* An observation map r_{q_θ} for this task discards exactly the θ-independent factor $h(x)$ and retains exactly the θ-dependent shape $g(\cdot,\theta)$.

Proposition 12.5. *Under this identification, T is sufficient for θ in the classical sense if and only if T is q_θ-sufficient in the sense of Theorem 10.2.*

Proof. By Theorem 12.4, T is classically sufficient exactly when $f(x;\theta) = g(T(x),\theta)h(x)$ for some θ-independent h, that is, exactly when the θ-dependent part of $f(\cdot;\theta)$ is a function of $T(x)$ alone. This is precisely the condition that $r_{q_\theta}(\omega(x))$, the retained θ-dependent shape at x, factors through T: two samples x_1, x_2 with $T(x_1) = T(x_2)$ necessarily have the same θ-dependent likelihood shape $g(T(x_1),\theta) = g(T(x_2),\theta)$ for every θ, which is exactly $r_{q_\theta}(\omega(x_1)) = r_{q_\theta}(\omega(x_2))$. By Theorem 10.2, this is exactly q_θ-sufficiency of T. □

This is a genuine specialization, not a loose analogy dressed up in new notation: the century-old statistical notion of sufficiency is the instance of this book's general sufficiency criterion obtained by setting q to be, specifically, the task of recovering how the data's likelihood depends on an unknown parameter. Every qualification this book has attached to q-sufficiency in general therefore applies to classical statistical sufficiency without modification, including Remark 11.1: a statistic can be highly non-injective, collapsing an enormous sample down to a single number, and still be sufficient, precisely because ω, here the θ-dependent likelihood shape, happens to be constant across everything the statistic merges.

Example 12.6. *Let $X_1, \dots, X_n$ be independent, identically distributed as $N(\mu,\sigma^2)$ with σ^2 known and μ the parameter of interest. The joint density is*

$$f(x;\mu) = (2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}\sum_i (x_i-\mu)^2\right].$$

Expanding $\sum_i (x_i-\mu)^2 = \sum_i x_i^2 - 2n\mu\bar{x} + n\mu^2$, exactly the trinomial-type expansion of Chapter 1 applied to n terms, gives

$$f(x;\mu) = \underbrace{(2\pi\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2}\sum_i x_i^2\right]}_{h(x)} \cdot \underbrace{\exp\left[\frac{n(2\mu\bar{x}-\mu^2)}{2\sigma^2}\right]}_{g(\bar{x},\mu)},$$

where $\bar{x} = \frac{1}{n}\sum_i x_i$ is the sample mean. Since g depends on x only through $\bar{x}$ and h does not depend on μ, Theorem 12.4 shows $T(x) = \bar{x}$ is sufficient for μ. By Proposition 12.5, the sample mean, a transformation collapsing n numbers down to one, is q_μ-sufficient: it discards everything about the individual x_i except their average, and nothing about the average is needed to recover it.

The sample mean is exactly the kind of non-injective transformation Remark 11.1 was written to keep in view. It merges an enormous number of distinct samples (any two samples with the same average are merged), and it is still fully sufficient, not despite being non-injective but consistently with the general theorem: the merged fibers of $\bar{x}$ are exactly the sets on which the μ-dependent part of the likelihood is constant, which is what the factorization computation verified directly.

12.5 Closing the substrate

This chapter has treated probability as the translation table of Chapter 11 promised: the same theorem, the same hierarchy of re-description, projection, and task-relative Collapse, instantiated with samples as histories and statistics as transformations. Variance and covariance gave the most direct possible restatement of Chapter 4's quadratic-form machinery; the parabola example gave a concrete, fully worked instance of Chapter 5's warning about vanishing corrections; portfolio risk showed the binding correction appearing with a negative sign in a setting where the sign has genuine economic content; and sufficiency showed that this book's general criterion is not merely compatible with an established statistical concept but literally identical to it, once the task is named precisely.

Chapter 13 turns to information theory, where the transformation of interest is a channel rather than a statistic, and where the relevant notion of what a transformation retains, mutual information, will turn out to measure something related to, but not identical with, the covariance-based coupling this chapter has just characterized.

Exercises

1. Compute $\mathrm{Var}(X_1 + X_2 + X_3)$ directly from Proposition 12.1 for three variables with $\mathrm{Var}(X_i) = 1$ for all i and $\mathrm{Cov}(X_i, X_j) = -0.4$ for all $i \neq j$, and confirm the result is smaller than $\sum_i \mathrm{Var}(X_i) = 3$.

2. Construct a second example, distinct from Example 12.2, of two random variables with zero covariance but a strong deterministic or near-deterministic relationship, and identify exactly which symmetry in your example forces the linear correlation to vanish.

3. For the two-asset portfolio with weights w_1, w_2 (with $w_1 + w_2 = 1$), variances $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = 1$, and $\mathrm{Cov}(X_1, X_2) = -1$ (perfectly negatively correlated, unit variance), find the weights that reduce σ_P^2 to exactly zero, and interpret the result in light of Proposition 12.1.

4. Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(θ). Using Theorem 12.4, show that $T(x) = \sum_i x_i$ is sufficient for θ, and explain, in the vocabulary of this chapter, what task T is q_θ-sufficient for and what specific information about the individual x_i it discards in the process.

Chapter 13

Entropy, Mutual Information, and Channels

13.1 Entropy and the chain rule

Let X be a discrete random variable with distribution $p(x)$. Its *entropy* is

$$H(X) = -\sum_x p(x) \log p(x),$$

a measure, in whatever base the logarithm is taken (bits, for base 2), of the average uncertainty in X before it is observed. For a pair (X, Y), the *joint entropy* $H(X, Y)$ is defined the same way over the joint distribution $p(x, y)$, and the *conditional entropy* is

$$H(X \mid Y) = \sum_y p(y)\, H(X \mid Y{=}y) = -\sum_{x,y} p(x, y) \log p(x \mid y).$$

Proposition 13.1 (Chain rule for entropy)**.** $H(X, Y) = H(Y) + H(X \mid Y)$.

Proof.

$$H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y) = -\sum_{x,y} p(x,y) \log\big[p(y)p(x \mid y)\big] = -\sum_{x,y} p(x,y) \log p(y) - \sum_{x,y} p(x,y) \log p(x \mid y$$

The first sum, marginalizing over x, is $-\sum_y p(y) \log p(y) = H(Y)$; the second is $H(X \mid Y)$ by definition. □

13.2 Mutual information as a relational correction

Definition 13.2. *The* mutual information *between X and Y is*

$$I(X;Y) = H(X) + H(Y) - H(X,Y).$$

By Proposition 13.1, $I(X;Y) = H(X)+H(Y)-[H(Y)+H(X \mid Y)] = H(X)-H(X \mid Y)$, and symmetrically $I(X;Y) = H(Y)-H(Y \mid X)$: mutual information is the reduction in uncertainty about X once Y is known, and, by the symmetry of the defining formula, exactly the same quantity is the reduction in uncertainty about Y once X is known.

The algebraic shape of $I(X;Y)$ is worth comparing directly to Δ_Q from Chapter 2. Define $\Delta_H(X,Y) = H(X,Y) - H(X) - H(Y)$, the entropy of the joint minus the sum of the marginal entropies; this is formally the binding correction of Chapter 2, with ordinary addition of numbers replaced by formation of a joint distribution. By definition, $\Delta_H(X,Y) = -I(X;Y)$. It is a standard fact, following from the non-negativity of Kullback–Leibler divergence (stated here without proof, in keeping with this book's convention for standard external results), that $I(X;Y) \geq 0$ always, with equality exactly when X and Y are independent. Consequently $\Delta_H(X,Y) \leq 0$ always: joint entropy is never more than the sum of the marginals, only ever equal to it (under independence) or strictly less.

This is a genuine and honest point of disanalogy with the quadratic case, not a detail to smooth over. $\Delta_Q(x,y) = 2xy$ in Chapter 2 could be positive, negative, or zero depending on the signs of x and y; the corresponding correction in this substrate, $\Delta_H(X,Y) = -I(X;Y)$, has a fixed sign, always non-positive, regardless of anything about X or Y beyond their both being well-defined random variables. The two substrates share the same algebraic shape, a joint evaluation minus a sum of separate evaluations, but the entropy instance of that shape happens to obey a sign constraint the quadratic instance does not. Recognizing this is exactly the discipline Chapter 2 asked for from the start: the shape of a binding correction is general, but its sign and range are properties of the specific evaluation, to be derived or cited, never assumed by analogy with a previously worked example.

13.3 Revisiting the parabola

Chapter 12's Example 12.2 showed that $\mathrm{Cov}(X,Y) = 0$ for $Y = X^2$ under a symmetric distribution on X, despite Y being a deterministic, and therefore maximally dependent, function of X. Mutual information, unlike covariance, is not restricted to detecting linear association, and it is worth confirming this directly rather than asserting it.

Example 13.3. *Let X be uniform on $\{-1,0,1\}$, each with probability $\frac{1}{3}$, and let $Y = X^2$, so $Y = 1$ with probability $\frac{2}{3}$ (when $X = \pm 1$) and $Y = 0$ with probability $\frac{1}{3}$ (when $X = 0$). As in the continuous case, $\mathbb{E}[X] = 0$ and $\mathbb{E}[XY] = \mathbb{E}[X^3] = \frac{1}{3}[(-1)^3 + 0^3 + 1^3] = 0$, so $\mathrm{Cov}(X,Y) = 0$. But Y is a deterministic function of X, so $H(Y \mid X) = 0$, and therefore*

$$I(X;Y) = H(Y) - H(Y \mid X) = H(Y) = -\left[\tfrac{1}{3}\log_2 \tfrac{1}{3} + \tfrac{2}{3}\log_2 \tfrac{2}{3}\right] \approx 0.918 \text{ bits},$$

strictly positive, despite $\mathrm{Cov}(X,Y) = 0$.

This is the concrete payoff of introducing a second substrate rather than working entirely within the quadratic-form apparatus of Part II. The covariance-based evaluation Q of Chapter 12 is, by construction, blind to this dependence, exactly as Chapter 5 warned any specific Q might be; mutual information is a different evaluation, sensitive to a broader class of dependence, and it registers the same underlying relationship as strictly positive. Nothing about this should be read as mutual information being a strictly superior or "more correct" measure of dependence in general; it is a different evaluation, with its own blind spots (Section 13.5 below exhibits one), and the moral of Chapter 5 continues to apply to it without exception: a vanishing $I(X;Y)$ certifies independence exactly, by the fact cited above, but a specific nonzero or zero value of any single evaluation should never be mistaken for a complete characterization of "how related" two quantities are.

13.4 Channels and the data processing inequality

A *channel* is a (possibly random) transformation from a source X to a received signal Y; in this book's vocabulary, it is a transformation T in exactly Chapter 10's sense, with the source's full realization playing the role of the history and the channel output playing the role of $T(H)$. A *decoder*, or any downstream task built on the channel's output, is exactly a task q in this book's sense, and the entire question of channel design, how much can a channel discard while still letting a decoder do its job, is a direct instance of the sufficiency question this book has been asking of every substrate so far.

Mutual information supplies a sharp form of this question when the channel output is processed further. Suppose $X \to Y \to Z$ forms a Markov chain, meaning Z depends on X only through Y: once Y is known, Z carries no further information about X beyond what Y already carries, formally $I(X; Z \mid Y) = 0$.

Theorem 13.4 (Data processing inequality, stated without proof)**.** *If $X \to Y \to Z$ is a Markov chain, then*

$$I(X; Z) \leq I(X; Y).$$

This is the information-theoretic form of a claim this book has made informally since Chapter 6: further processing of an already-collapsed representation cannot recover distinctions that processing has discarded, and, in this precise quantitative sense, it cannot even partially recover them, only preserve or further reduce what remains. If $Z = T_2(Y)$ for some further transformation T_2 applied to a channel output $Y = T_1(X)$, Theorem 13.4 says the composite $T_2 \circ T_1$ can never carry more information about the original source X than T_1 alone did. In the vocabulary of Chapter 10's hierarchy, composing transformations moves a representation further into, or leaves it at the same depth within, the region of possible task-relevant loss relative to the original history; it never moves it back out. Equality in Theorem 13.4 is possible, exactly when T_2 discards nothing further that mattered for recovering information about X, which is the information-theoretic counterpart of Proposition 10.3: a further transformation that happens to be a re-description relative to the task "preserve information about X" leaves the inequality tight.

13.5 Synergy: when the whole exceeds the sum, exactly

Chapter 2 left an exercise unresolved, later revisited in Chapter 11's closing section: the additivity of pairwise corrections that holds for $Q(x) = x^2$ fails for $Q(x) = x^3$, whose expansion contains an irreducible three-way term, $6xyz$, and Chapter 11 identified the graph-theoretic name for this phenomenon as a hyperedge. Information theory supplies a clean, fully computable instance of the same phenomenon.

Example 13.5 (Synergy)**.** *Let X and Y be independent fair bits, and let $Z = X \oplus Y$ (exclusive or). Since Z is a fixed function of the independent pair (X, Y) with Z itself uniform on $\{0, 1\}$, and since Z is independent of X alone (knowing X alone gives no information about Z, since Y is independent and unknown) and independent of Y alone by the same argument,*

$$I(X; Z) = 0, \qquad I(Y; Z) = 0, \qquad \textit{so} \qquad I(X; Z) + I(Y; Z) = 0.$$

But Z is a deterministic function of the pair (X, Y) together, so $H(Z \mid X, Y) = 0$ and

$$I(X, Y; Z) = H(Z) - H(Z \mid X, Y) = H(Z) = 1 \textit{ bit.}$$

Hence $I(X, Y; Z) = 1 > 0 = I(X; Z) + I(Y; Z)$.

This is *synergy* in the precise technical sense this book has been building toward since Chapter 2's cubic exercise: information about Z that is available from the pair (X, Y) jointly and from neither X nor Y separately, in any amount whatsoever. It is the information-theoretic realization of an irreducible three-way relation, exactly as $6xyz$ was the algebraic realization and a hyperedge was the graph-theoretic realization in Chapter 11. No pairwise summary of how X relates to Z and how Y relates to Z, however finely measured, can recover this fact; the joint pair has to be considered as a single, addressable unit for the dependency to become visible at all, which is the same lesson Chapter 11 stated for hypergraphs, now demonstrated with a fully worked, three-variable example rather than left as a diagnostic to recognize.

13.6 Closing the substrate

Entropy and mutual information have completed the translation promised in Chapter 11: the same binding-correction shape as Chapters 2 and 4, here obeying a fixed sign that the quadratic case did not; the same vanishing-correction caution as Chapter 5, here resolved concretely by an evaluation with a wider detection range, itself subject to the identical caution in turn; and, with the data processing inequality, a fully rigorous, quantitative form of this book's recurring claim that processing a representation can only preserve or reduce what remains recoverable from it, never restore what an earlier step has already discarded. Chapter 14 turns from these largely static, single-evaluation questions to the case where history accumulates over time, and where the order in which bindings occur, not merely which bindings occur, becomes part of what a later evaluation needs to reconstruct.

Exercises

1. Compute $H(X)$ for the three-valued X in Example 13.3, and confirm directly that $I(X; Y) = H(X) + H(Y) - H(X, Y)$ agrees with the value computed in the text via $H(Y) - H(Y \mid X)$.

2. Construct a second example of two variables with zero mutual information but a nonzero higher-order statistical relationship of your choosing, or explain why $I(X; Y) = 0$ is a stronger, more restrictive condition than $\mathrm{Cov}(X, Y) = 0$ and no such example can exist within the same class of dependence covariance already ruled out.

3. Using the chain rule for entropy applied twice, to $H(X, Y, Z)$ expanded first as $H(X) + H(Y, Z \mid X)$ and again as $H(Y) + H(X, Z \mid Y)$, derive the chain rule for mutual information, $I(X; Y, Z) = I(X; Y) + I(X; Z \mid Y)$, and use it together with the non-negativity of conditional mutual information to reproduce the proof sketch of Theorem 13.4.

4. In Example 13.5, compute $I(X; Y \mid Z)$ and confirm it is strictly positive, despite X and Y being independent unconditionally. Explain, in the vocabulary of this chapter, why conditioning on Z can introduce an association between two variables that were independent before conditioning.

Chapter 14

Path Dependence, Hysteresis, and Repair

14.1 Path-sensitive evaluations, redeemed

Chapter 9 flagged, and explicitly deferred, a second reason Proposition 9.1 can fail: a state map σ that depends on the sequence of intermediate states a world passed through, not merely on the final extensional structure $\tau(H) = (P(H), R(H), B(H))$. This chapter takes up that case directly.

Definition 14.1. *A state map σ is* path-sensitive *if it does not factor through τ: there exist well-formed histories H, H' with $\tau(H) = \tau(H')$ but $\sigma(H) \neq \sigma(H')$.*

Path sensitivity is not a defect in σ; it is simply a fact about which evaluations are of interest in a given substrate. Every evaluation this book has used through Chapter 13, quadratic forms, covariances, entropies, has been extensional in Chapter 9's sense, depending only on final structure. Many evaluations of genuine interest are not, and the clearest examples come from physical systems whose present behavior depends on their history of loading, not merely their present state.

14.2 Hysteresis

Example 14.2 (Magnetic hysteresis)**.** *Let H record a sequence of applied external field values, $h_1, h_2, \ldots, h_t$, and let $\sigma(H)$ denote the resulting magnetization of a ferromagnetic material after that sequence has been applied. Let $\tau(H)$ record only the current field value, h_t, discarding the sequence that led to it. Two histories H and H' can share $\tau(H) = \tau(H') = h_t$, the field currently sits at the same value in both, while differing in whether h_t was reached by increasing the field from a lower value or decreasing it from a higher one. Ferromagnetic materials are exactly the physical systems for which $\sigma(H) \neq \sigma(H')$ in this situation: the same current field corresponds to two different magnetizations, depending on the direction of approach, the phenomenon named* hysteresis*.*

Hysteresis is path sensitivity with a name and a long empirical history, and Example 14.2 is a case where τ, the current field value alone, is too coarse a summary for any σ of practical interest: no evaluation of magnetization can be extensional in Chapter 9's sense if it is required

to depend only on the current field, because the physical system itself does not behave that way. This is worth stating as a general moral before the machinery of this chapter is applied more abstractly: sometimes the failure of Proposition 9.1's hypothesis is not a shortcoming of a poorly chosen σ, but a correct diagnosis that the substrate genuinely requires more of τ than the final popped, refused, and bound sets, up to and including the order or direction in which they were assembled.

14.3 Irreversibility of an append-only history

Chapter 8 established that H is append-only: no operation in this book's calculus edits or deletes an entry once recorded. This has a consequence for what "undoing" an operation can and cannot achieve, worth stating precisely.

Proposition 14.3. *Let H_{t_0} be a history at time t_0, and let H extend H_{t_0} by some further operations (an "excursion") followed by an operation designed to restore the extensional structure to its value at t_0, so that $\tau(H) = \tau(H_{t_0})$. Then, in general, $\omega(H) \neq \omega(H_{t_0})$.*

Proof. By example. Suppose a world's governing admissibility rule for Bind specifies that an entry which has ever appeared in $B(H)$, even if a later operation removes that specific pair from the currently active bound set, may not be freshly bound to a third entry without an explicit intervening Refuse-clearing step, a rule no less legitimate than any other admissibility constraint this book has considered. Let H_{t_0} pop a and b with no binding between them. Let H extend H_{t_0} by binding a to b, then recording a further operation that removes this pair from the currently active bound set (restoring $\tau(H)$'s bound-pair component to match $\tau(H_{t_0})$, since the pair is no longer active in either). At H_{t_0}, $\omega(H_{t_0})$ permits binding a freely to a new entry c. At H, by the stated rule, $\omega(H)$ does not permit this without the clearing step, since a's history includes having once appeared in $B(H)$. Hence $\omega(H) \neq \omega(H_{t_0})$ despite $\tau(H) = \tau(H_{t_0})$. □

The content of Proposition 14.3 is not that reversal is impossible in some vague metaphysical sense; it is the precise, structural observation that H, being append-only, retains the record of the excursion permanently, and any admissibility rule sensitive to that record, not merely to the currently active extensional summary $\tau(H)$, can leave $\omega(H)$ different from what it was before the excursion, even after the excursion's visible effects on τ have been nominally undone. This is the operational, calculus-level counterpart of Example 14.2: hysteresis in a physical magnet and the admissibility-rule example here are the same structural phenomenon, a governing rule (physical law in one case, a world's stipulated admissibility rule in the other) that consults more of H than its coarse current summary $\tau(H)$ reveals.

14.4 Repair as relational reconstruction

The word "repair" has been used informally throughout this book (Chapter 3's audit concerns, Chapter 8's brief mention of repair as a downstream task Refuse and Collapse must remain compatible with) without a working definition. This section supplies one, and uses it to settle a question Chapter 9's "binding before revision" section left as a caution rather than a resolved method: given that a popped value was wrong, what does it actually take to fix the resulting evaluation correctly.

Definition 14.4. *A* repair *of an entry $a \in E(H)$, discovered to carry an erroneous value, is a sequence of operations that pops a corrected entry a' and re-establishes, via new* Bind *operations, whatever relational structure a previously participated in, using a' in place of a.*

The definition is deliberately restrictive: it does not permit editing a's recorded value directly, since H is append-only, and it does not permit patching a collapsed scalar output directly either, requiring instead that the repair operate on the relational structure, $B(H)$, itself.

Proposition 14.5 (Patching is not repair)**.** *Let H contain* Bind(a, b) *under the trinomial evaluation Q, with a's value later discovered to be erroneous, and let $s = \sigma(H) = Q(a + b)$ be the collapsed scalar. Suppose a proposed "repair" instead adjusts s directly by an amount δ computed from the difference between a's erroneous and corrected values, without popping a' or re-invoking* Bind*. Then this adjustment reproduces the correctly repaired value $Q(a' + b)$ only when δ is chosen to exactly equal $Q(a' + b) - Q(a + b)$, which requires knowing b specifically, exactly the information the direct-patch approach was intended to avoid consulting.*

Proof. $Q(a' + b) - Q(a + b) = (a')^2 + b^2 + 2a'b - a^2 - b^2 - 2ab = (a')^2 - a^2 + 2b(a' - a)$, which depends on b explicitly (through the cross term $2b(a' - a)$), not merely on a and a'. A patch computed from a and a' alone, without consulting b, cannot in general produce this value. □

This is Chapter 9's "binding before revision" result, now given its proper name and its constructive resolution. A repair that only touches the collapsed scalar, without re-consulting the entries the erroneous value was bound to, is not a repair in the sense this section defines; it is a patch, and Proposition 14.5 shows precisely why it generally fails: the correction to a bound evaluation depends on the specific coupling, not on the erroneous and corrected values alone. A repair that pops a' and re-binds it to b, recomputing $Q(a' + b)$ from the actual current value of b, succeeds trivially, by construction, because it reconstructs the relation rather than attempting to infer its effect from outside.

This also gives a precise, constructive answer to the question Chapter 10 left open in a different but related form: which evaluations support principled repair at all. An evaluation σ_B that reads $B(H)$ directly, as in Proposition 10.7, exposes exactly the relational structure a repair needs to re-traverse; an evaluation T_Δ that reads only popped values, blind to $B(H)$, gives a repairer no way to know which pairs need re-binding at all, since it never distinguished bound pairs from merely co-occurring ones in the first place. Whether a system supports relational repair, in the sense of this section, and whether its evaluations can serve as a witness of binding, in the sense of Chapter 10, turn out to depend on the same underlying fact: whether the evaluation in use consults the operational record $B(H)$ or only the values that record happens to be attached to.

14.5 Closing Part IV

Part IV set out to show that the general theory of Parts I through III is not a free-floating formalism but a genuine spine running through concrete substrates. Graphs gave the theory its most direct structural instance, with the binding matrix of Chapter 4 read as an adjacency matrix and quotients read as transformations subject to Theorem 10.2. Probability gave the theory its sharpest correspondence with an existing, independently developed vocabulary, classical statistical sufficiency turning out to be an exact instance of q-sufficiency once the task is named correctly. Information theory gave the theory its most quantitatively rigorous form, the data

processing inequality standing as a precisely stated, if externally cited rather than internally derived, version of this book's recurring claim that processing cannot manufacture what an earlier step has discarded, while also supplying, in the synergy example, a fully worked instance of the irreducible higher-order relation Chapter 2 first flagged and could not, at that point, resolve. And this chapter has shown that the same apparatus, applied to time-indexed histories, gives hysteresis, irreversibility, and principled repair their formal homes, closing the loop on Chapter 9's deferred path-sensitivity case and on Chapter 9's own "binding before revision" caution, this time with a constructive resolution rather than only a warning.

Part V turns from these substrate-by-substrate instances to the question that has been implicit in every one of them: what, in general, makes a Collapse admissible, and what is actually recoverable once one has occurred.

Exercises

1. Construct a second physical or economic example of path sensitivity, distinct from magnetic hysteresis, and identify explicitly what τ would need to include, beyond the current coarse summary, for the relevant σ to become extensional.

2. In Proposition 14.3's proof, describe what an explicit "Refuse-clearing" operation would need to record in order to restore $\omega(H)$ to $\omega(H_{t_0})$ exactly, and confirm that adding such an operation does not violate H's append-only property.

3. Extend Proposition 14.5 to a case with three entries a, b, c all pairwise bound under the trinomial evaluation, where a's value is corrected. Determine which pairwise terms in the collapsed scalar must be recomputed, and confirm that a patch computed from a and a' alone cannot reproduce the correct total without consulting both b and c.

4. Using the distinction between σ_B and T_Δ from Chapter 10, explain in your own words why a system logging only final aggregate totals (analogous to T_Δ) is structurally unrepairable in the sense of this chapter, even if every individual erroneous value is later identified and corrected in isolation.

Part IV

Admissibility and Collapse

Chapter 15

The Interaction Ledger

15.1 What Parts II through IV have all been building

Every substrate visited in Part IV kept the same four kinds of information, under different names. A quadratic form (Chapter 4) had diagonal self-terms a_{ii} and off-diagonal couplings a_{ij}. A graph (Chapter 11) had vertex weights and edge weights. A random vector (Chapter 12) had individual variances $\mathrm{Var}(X_i)$ and pairwise covariances $\mathrm{Cov}(X_i, X_j)$. A collection of signals (Chapter 13) had individual entropies $H(X_i)$ and pairwise mutual informations $I(X_i; X_j)$. In every case, alongside these two kinds of numerical content, there was also an operational history, Chapter 7's H, recording how the entries came to be popped and bound, and a set of admissibility constraints, Chapter 8's R_H, recording which identifications are forbidden regardless of what any evaluation computes. This chapter gives this four-part structure a single name and treats it as the object whose fate any Collapse operation actually decides.

Definition 15.1. *The* interaction ledger *of a world with popped entries $x_1, \ldots, x_n$ under evaluation Q is the tuple*

$$L = \big(\{U_i\},\, \{R_{ij}\},\, H,\, A\big),$$

where U_i is the unary self-contribution of x_i (a diagonal entry a_{ii}, a variance $\mathrm{Var}(X_i)$, an entropy $H(X_i)$, according to substrate), R_{ij} is the relational contribution between x_i and x_j for $i < j$ (an off-diagonal a_{ij}, a covariance, a mutual information), H is the full operational history in Chapter 7's sense, and A is the admissibility data, the refusal set $R_H \subseteq E(H) \times E(H)$ together with any further declared admissibility rules a given world imposes.

Nothing in this definition is new content; it is a relabeling that makes visible a pattern that has been present, unremarked, since Chapter 4. What is new is the observation that every Collapse this book has performed can be classified by exactly which pieces of L it retains and which it discards, and that classification, not the specific numerical formula involved, is what determines a Collapse's admissibility for a given task.

15.2 A taxonomy of collapses

Let $S \subseteq \{U, R, H, A\}$ denote the components of L a given Collapse retains, writing U for the full unary profile $\{U_i\}$, R for the full relational profile $\{R_{ij}\}$, and similarly H, A as before. Four points in this space are worth naming, since they recur constantly in ordinary practice, in this book, and elsewhere.

Total-only ($S = \varnothing$, retaining only the scalar $\sum_i U_i + 2\sum_{i<j} R_{ij}$ itself, not even the individual U_i or R_{ij} separately): the most aggressive collapse this book has considered, exemplified by the trinomial square reduced to a single number, or T_Δ from Chapter 10 read at a single pair.

Profile-only, unary ($S = \{U\}$): retaining each U_i individually but discarding every R_{ij}. This is an extremely common real-world collapse, worth naming because it is easy to perform without noticing: reporting each asset's individual variance while discarding the covariance matrix, reporting each source's entropy while discarding pairwise mutual information, reporting each vertex's total degree while discarding which specific edges contributed to it. It is also, by Proposition 12.1 and its analogues, provably insufficient for any task requiring the true joint evaluation, since the joint total cannot be recovered from the unary profile alone whenever any $R_{ij} \neq 0$.

Extensional ($S = \{U, R, A\}$, retaining full unary and relational profiles and admissibility data, but discarding H itself): this is exactly Chapter 9's $\tau(H) = (P(H), R(H), B(H))$, now seen to be a specific, named point in the ledger taxonomy rather than an ad hoc construction local to that chapter. Proposition 9.1's hypothesis, that σ factors through τ, is precisely the hypothesis that a task never needs more of L than this extensional point provides.

Full ledger ($S = \{U, R, H, A\}$): retaining everything, which, since U and R are themselves computable from H given the evaluation Q, is informationally equivalent to retaining H and A alone. This is not really a Collapse at all in the sense of Chapter 10; it is closer to a re-description, since nothing has been merged or discarded.

15.3 The general admissibility theorem, specialized to ledgers

Theorem 15.2 (Ledger sufficiency)**.** *Let T_S denote the transformation retaining exactly the ledger components in $S \subseteq \{U, R, H, A\}$ and discarding the rest. Then T_S is q-sufficient, in the sense of Theorem 10.2, if and only if no two histories agreeing on every component in S disagree on $r_q(\omega(H))$.*

Proof. Immediate specialization of Theorem 10.2 to $T = T_S$. □

The value of Theorem 15.2 is not its proof, which is a one-line specialization, but the vocabulary it licenses: instead of asking, case by case, whether some particular collapse is admissible for some particular task, one can ask which named region of the ledger taxonomy a task's needs fall into, and read off admissibility from that classification directly.

Corollary 15.3. *The following identifications hold, each already established in earlier chapters and now unified under one criterion.*

1. *A task requiring the distinction between the generative history $xy + yx$ and the collected $2xy$ (Chapter 3) needs $H \in S$; retaining U, R, A alone, the extensional point, is provably insufficient for it, exactly as Chapter 3 argued directly.*

2. *A task requiring the specific edge weight between two named vertices after a graph quotient (Chapter 11's Example) needs $R \in S$; retaining only the aggregate total is insufficient, exactly as that example showed by direct construction.*

3. *A task requiring verification that a proposed identification does not violate a recorded refusal (Chapter 8) needs $A \in S$; discarding admissibility data entirely makes this question unanswerable regardless of what else is retained.*

4. *A task requiring only the total variance, entropy, or evaluated total of a system (most of Chapter 12's portfolio-risk use case) needs none of U, R, H, A individually; the total-only collapse is fully sufficient for it.*

Each of these four claims was proved, or argued directly, in its originating chapter, using whatever apparatus was locally available. Corollary 15.3 does not add new content to any of them; it records that they are four instances of one theorem, differing only in which single letter of $\{U, R, H, A\}$ the task in question happened to require, and that observation is what makes the ledger a genuine unification rather than a decorative relabeling.

15.4 Conservation requirements

The applications literature this book has been drawing on, audits, portfolio management, evidentiary standards, informally uses a wider vocabulary of things a summary might be required to "preserve": total value, sign, ordering, reachability, provenance, uncertainty, reversibility, future option structure. Each of these turns out to be a specific instance of q-sufficiency relative to a specific choice of S, rather than an independent requirement needing its own theory.

Preserving total value is total-only sufficiency, the weakest requirement in the taxonomy, needing no component of S beyond the scalar itself. *Preserving sign* is a coarsening of this, needing even less: only the sign of the total, not its magnitude. *Preserving ordering* among several systems, knowing which of two totals is larger without needing either exactly, is coarser still. *Preserving provenance*, in Chapter 3's sense, requires $H \in S$ specifically, nothing less. *Preserving reversibility*, the ability to reconstruct an earlier state after some operations, is exactly re-description in Chapter 10's sense, and Proposition 10.3 already characterized when it holds. *Preserving future option structure* is Chapter 7's Ω-sufficiency directly, the very criterion this whole apparatus was built from. None of these are separate demands on a theory of collapse; they are different named points along the single axis Theorem 15.2 already parametrizes, and a project that tries to satisfy "as many conservation requirements as possible" without naming which one a given task actually needs is asking an ill-posed question, for the same reason Chapter 8 first flagged: admissibility is only ever relative to a stated q, and q has to be named before "preserves enough" has a truth value at all.

15.5 Closing remarks

The interaction ledger is not a new theory. It is the vocabulary this book has been using since Chapter 4, made explicit enough to state once, generally, what had previously been re-derived separately for quadratic forms, graphs, probability, and information. Chapter 16 uses this vocabulary to give premature collapse, and Refuse's role in constraining it, their most general statement yet; Chapter 17 closes Part V by asking, systematically, what "recoverable" actually means once several different targets of recovery, H, R, A, or merely the total, are all on the table at once, rather than considered one at a time as the applications chapters necessarily did.

Exercises

1. For the sample-mean statistic of Chapter 12 (Example 12.6), classify which region of the $\{U, R, H, A\}$ taxonomy it occupies, treating the individual data points as the popped

entries and $\bar{x}$ as the retained summary.

2. Construct a task for which the profile-only unary collapse ($S = \{U\}$) is sufficient, and a second, closely related task for which it is not, in the style of Chapter 11's quotient example.

3. Reformulate Chapter 9's "binding before revision" example (Section 9.5) in ledger vocabulary: which component of L does a correct repair need to consult that a naive scalar patch does not?

4. Give an example of a real reporting practice (financial, scientific, or administrative) that performs a profile-only unary collapse without stating so explicitly, and identify a plausible downstream task for which that omission would matter.

Chapter 16

Premature Collapse and Refusal as a Constraint on Quotienting

16.1 Premature collapse, generalized

Chapter 7 defined premature collapse narrowly, in terms of a single scalar state map: a collapse via σ is premature exactly when $\sigma(H_1) = \sigma(H_2)$ but $\omega(H_1) \neq \omega(H_2)$. With the ledger of Chapter 15 in hand, this definition can be restated at the level of generality the rest of this book has been working toward.

Definition 16.1. *A collapse T_S, retaining ledger components $S \subseteq \{U, R, H, A\}$, is* premature relative to a task q *if T_S fails to be q-sufficient in the sense of Theorem 15.2: some pair of histories agreeing on every component in S disagree on $r_q(\omega(H))$.*

This is not a new claim; it is Chapter 7's original definition, specialized to $\sigma = T_\varnothing$, generalized outward to arbitrary S. What the generalization buys is precision about *which* discarded component was responsible: a collapse can be premature specifically because it discarded H (Chapter 3's provenance cases), specifically because it discarded R (Chapter 11's specific-edge-weight case), or specifically because it discarded A, a case this chapter now turns to, since it has not yet been given its own worked treatment.

16.2 Two independent grounds for blocking a collapse

Chapter 8 already distinguished two forms of inadmissibility for Collapse in general: task-relative inadmissibility, failing q-sufficiency for a stated task, and a stronger, unconditional inadmissibility, violating a recorded refusal regardless of what any task needs. It is worth stating this distinction one more time, now specifically for the quotienting operations of Chapter 11, since "constraint on quotienting" is precisely the setting in which the two grounds are most easily conflated in practice.

Theorem 16.2 (Two-fold admissibility of a quotient)**.** *Let T_S be a quotient merging some set of entries. T_S is a legitimate collapse if and only if both of the following hold:*

(a) Task-relative sufficiency*: T_S is q-sufficient for the task q at hand, in the sense of Theorem 15.2.*

(b) Refusal-respecting*: T_S does not identify any pair $(a, b) \in A$, that is, the quotient does not merge, aggregate, or otherwise treat as equivalent any pair the world's history has explicitly refused.*

Proof. Immediate from the definitions: (a) is q-sufficiency directly, and (b) is exactly Chapter 8's condition for a Collapse to avoid the stronger, task-independent form of inadmissibility, restated for a quotient-type T_S specifically, since a quotient's "identifying a and b" is exactly its $\sim_\sigma$-relation in Chapter 8's sense holding of that pair. □

The two conditions are independent, and it is worth having a worked instance of each failing without the other, since collapsing them into a single undifferentiated notion of "bad merge" is exactly the kind of looseness this book has tried to discipline throughout.

Example 16.3 (Condition (b) fails, (a) would have held)**.** *Recall Chapter 8's witnesses a, b, refused because both trace to a common source. Suppose the task at hand is merely "report the total number of testimony items received," a task for which merging a and b into a single count would, numerically, cause no problem at all: condition (a) would be satisfied, since the total count is recoverable either way. Condition (b) fails regardless: $(a, b) \in A$, and the quotient is blocked outright, independent of the fact that this particular task would have tolerated it. The refusal is a fact about the pair, not about any specific downstream use.*

Example 16.4 (Condition (a) fails, (b) holds)**.** *Return to Chapter 11's four-vertex quotient, merging vertices* 1 *and* 2 *with no refusal ever recorded between them, so condition (b) is satisfied trivially. For the task "report the edge weight between* 1 *and* 2 *specifically," condition (a) fails, exactly as shown there: the quotient discards R_{12}, and no refusal is involved in that failure at all. The two vertices were always free to be merged, in the sense that nothing in A forbade it; the merge is simply the wrong tool for that particular question.*

16.3 The missing-rectangle test

Theorem 16.2's condition (a) is, in general, only checkable once a specific task q has been named. It is worth having a coarser, task-independent diagnostic that flags a specific and common way collapses go wrong, even before any task has been fixed, precisely because it is the check most people would want to run first, informally, before committing to a merge.

Definition 16.5 (Missing-rectangle test)**.** *Given a proposed quotient merging a group of entries $G = \{x_{i_1}, \ldots, x_{i_k}\}$, compute R_{ij} for every pair (i, j) with at least one of i, j in G and the other outside it, or with both in G but not otherwise identified by the quotient. If any such $R_{ij} \neq 0$, the test flags the quotient.*

Proposition 16.6. *If the missing-rectangle test flags no nonzero R_{ij} for a proposed quotient, the quotient is total-value-conserving in the sense of Chapter 15: the collapsed total is recoverable as the sum of the quotient's own internal terms, with no correction needed for interactions crossing the merge boundary. If the test flags a nonzero R_{ij}, the quotient is not total-value-conserving without further adjustment.*

Proof. Immediate from the definition of the binding correction (Chapter 2) and its ledger form (Chapter 15): the total evaluation decomposes as $\sum_i U_i + 2\sum_{i<j} R_{ij}$, and any R_{ij} crossing the merge boundary that is discarded, rather than absorbed correctly into the new quotient

structure, is exactly the amount by which the naive quotient total will differ from the true total. □

The test is deliberately modest in what it certifies, and the modesty is the point, not a limitation to apologize for. A vanishing R_{ij} across every boundary pair certifies only *total-value conservation*, the weakest item in Chapter 15's conservation taxonomy. It says nothing about whether H or A are safe to discard, and Chapter 5's four-fold list of explanations for a vanishing correction applies here without any softening: a zero R_{ij} at the moment of testing could reflect genuine absence of coupling, a coupling this particular evaluation cannot detect, an exact cancellation, or measurement too coarse to see it. The missing-rectangle test is a fast, cheap, and useful first check; it is not, and was never intended to be, a substitute for Theorem 16.2's two actual conditions.

16.4 A worked synthesis

Bringing the pieces of this chapter together: suppose a graph quotient is proposed, merging vertices 3 and 4 of some larger network, for the task of estimating total flow between two named regions of the graph. The correct procedure, in light of this chapter, has three steps, not one. First, run the missing-rectangle test on the boundary of $\{3, 4\}$: if every crossing R_{ij} is zero, the merge is at least total-value-conserving, and no correction term needs to be carried forward. Second, check condition (b) of Theorem 16.2: has any refusal ever been recorded between 3 and 4, or between either of them and some other entry that this specific merge would inadvertently identify? If so, the merge is blocked outright, regardless of what the first step found. Third, and only once the first two checks are satisfied, check condition (a) directly against the actual task: does the flow-estimation task need anything beyond what the quotient retains, such as the specific internal composition of the merged pair, in some scenario the task's own definition covers? Only a merge that clears all three steps is a legitimate collapse for that task, and each step is doing work the other two cannot substitute for.

16.5 Closing remarks

This chapter has done two things with the ledger vocabulary introduced in Chapter 15. It has generalized premature collapse from a single scalar-based definition to a claim about which named ledger component a task actually needed and a collapse discarded, and it has shown, concretely for quotienting, that task-relative insufficiency and outright refusal-violation are independent failure modes, requiring independent checks, with the missing-rectangle test supplying a fast, honestly limited diagnostic for one specific and common kind of the first failure. Chapter 17 closes Part V by asking the question this chapter has assumed an answer to at every step, what it actually means for something to be *recoverable* at all, now that several distinct targets of recovery, H, R, A, and the various coarsenings of each, are all visibly on the table rather than considered one at a time.

Exercises

1. Construct a five-vertex graph and a proposed three-vertex merge for which the missing-rectangle test passes (all boundary $R_{ij} = 0$) but for which the merge is nonetheless blocked

by a recorded refusal, illustrating that passing the test does not imply condition (b).

2. Construct a second example, distinct from Chapter 11's, in which condition (a) fails for a specific task despite the missing-rectangle test passing, showing that total-value conservation does not imply task-sufficiency for tasks that need something other than the total.

3. Explain why the missing-rectangle test, applied to the parabola example of Chapters 12 and 13 ($Y = X^2$), would not by itself have revealed the dependence that mutual information detected, and identify precisely which of Chapter 5's four explanations for a vanishing correction is responsible.

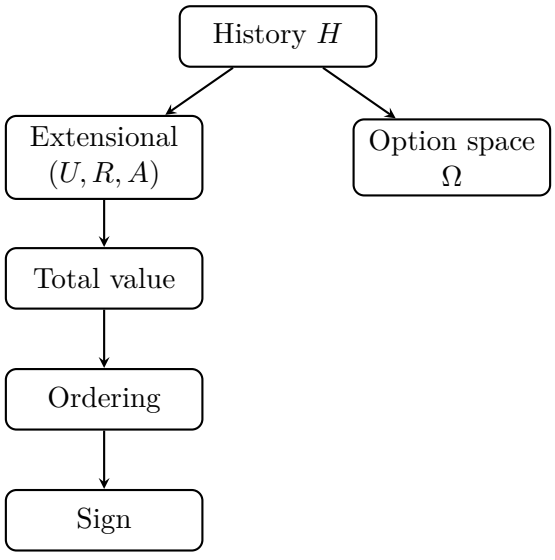
4. Using Theorem 16.2, describe a real administrative or organizational scenario in which a proposed merger of two records, teams, or accounts should be blocked for reason (b) even though it would clearly satisfy reason (a) for the specific report currently being requested.

Chapter 17

Recoverability

17.1 A lattice, not a chain

It is tempting, having built a single taxonomy of ledger components in Chapter 15, to treat recoverability as one scale: more is preserved, or less, and any two collapses can be ranked against each other by how much of the ledger survives. This chapter's main claim is that this temptation should be resisted. Some recoverability targets do form a genuine chain, each strictly implying the ones below it, but at least two targets of real interest are *incomparable*: neither implies the other, and a collapse can achieve one while failing the other completely. Confusing a lattice for a chain is exactly the kind of oversimplification Chapter 10 warned against when it insisted that non-injectivity alone does not determine a collapse's severity; this chapter gives that warning its full, structured form.



Each arrow denotes strict logical implication: recovering the target at the tail of an arrow guarantees recovering the target at its head, but not conversely, and the absence of an arrow between Extensional and Ω is deliberate, recording that neither recoverability target implies the other.

17.2 The main chain

Proposition 17.1. *History recovery implies extensional recovery; extensional recovery implies total-value recovery; total-value recovery implies ordering recovery (the ability to compare two totals without knowing either exactly); and ordering recovery implies sign recovery (the ability to determine whether a total is positive, negative, or zero).*

Proof. Each implication is immediate from the definitions already established. Given H, the unary and relational profiles U, R are computable directly from the popped values under the governing evaluation, and A is simply read off H as its refusal component; this is Chapter 15's observation that the full ledger is informationally equivalent to $\{H, A\}$. Given the extensional profile, the total value $\sum_i U_i + 2\sum_{i<j} R_{ij}$ is computed by the stated formula. Given two exact totals, their relative order is determined trivially by comparing them. Given the relative order of a total against zero specifically, its sign is determined trivially. □

Each step in this chain is a genuine collapse, in Chapter 10's sense, discarding strictly more than the step before it, and each is total-value-conserving, in Chapter 15's vocabulary, only up to the point in the chain where total value itself is retained; ordering and sign recovery are strictly coarser than total-value recovery; and Proposition 17.1 does not claim, and should not be read as claiming, that any of these coarser targets is somehow "almost as good" as the finer ones above it. A task requiring the exact total gains nothing from a collapse that only guarantees ordering, however far down the chain that collapse otherwise sits.

17.3 Two incomparable targets

The chain above accounts for everything below Extensional, but H has a second child in the diagram, Ω, Chapter 7's option space, and the two are not comparable.

Proposition 17.2. *Extensional recovery does not imply Ω-recovery, and Ω-recovery does not imply extensional recovery.*

Proof. For the first direction: consider Chapter 14's magnetic hysteresis example. Two histories H, H' can share identical current field value and identical current magnetization, hence identical extensional data under any evaluation built from these two quantities alone, while differing in whether the field was most recently increasing or decreasing. The admissible future responses to a further small perturbation, Ω, differ between the two: a material on the ascending branch and one on the descending branch respond differently to the same next increment, by the defining nature of hysteresis. Hence extensional recovery (of field and magnetization) holds while Ω-recovery fails.

For the second direction: consider a world whose admissibility rule for future Pop and Bind operations depends only on which entries have been popped and bound so far, structurally, and not on their specific values, so that $\Omega = \omega(H)$ is fully determined by the structural pattern of H alone. Let $T(H) = \omega(H)$ directly, so Ω-recovery holds trivially. Two histories with the same structural pattern but different popped values (different U_i) share the same Ω under this rule, yet differ in their extensional data. Hence Ω-recovery holds while extensional recovery fails. □

The two counterexamples are worth holding side by side, since they fail for opposite reasons. The hysteresis case fails because Ω depends on something (the direction of approach) that the

extensional summary, built only from current values, never captured in the first place; more of H was needed than the extensional point provides, exactly Chapter 14's diagnosis of path sensitivity restated in this chapter's vocabulary. The structural-rule case fails because Ω, in that particular world, happens not to depend on the specific values at all, so knowing Ω exactly tells a reader nothing about U or R, which live in a part of the ledger Ω was never built to track. Neither case is exotic or contrived beyond what is needed to make the logical point; each corresponds to a real distinction, between a system where the future depends on the path taken and a system where the future depends only on which distinctions exist, not their values, and both kinds of system are common enough that assuming either implication holds by default would be a real error, not merely a theoretical gap.

17.4 Weak recoverability is not weak reconstruction

The chain of Proposition 17.1, and the incomparability of Proposition 17.2, together give this chapter's title its full meaning. "Recoverable" is not a single yes-or-no property, nor even a single graded property admitting a unique measure of "how much." It is a family of distinct properties, each with its own precise definition, related to one another by a partial order that is genuinely a lattice, with a top (H) and a strict chain of increasingly coarse totals beneath one branch, but with two incomparable second-tier nodes rather than a single linear ranking. A report that says a collapse "preserves most of the important information" has not yet said anything checkable until it specifies which node, or nodes, of this lattice it is claiming to sit at, and Chapter 16's two-fold admissibility theorem already showed, in a narrower setting, exactly why that specification cannot be skipped: task-relative sufficiency and refusal-respect were shown there to be independent conditions, and this chapter's incomparability result is the same discipline applied one level higher, to the recovery targets themselves rather than to the tasks that call on them.

17.5 Closing Part V

Part V set out to answer, in general, the question every substrate-specific chapter in Part IV had answered locally: what does it mean for a collapse to be admissible, and what does it actually preserve. Chapter 15 gave every substrate's separate self-term/cross-term/history/admissibility structure one common name, the interaction ledger, and showed that four previously separate results were instances of one theorem. Chapter 16 used that vocabulary to give premature collapse its most general statement and to separate task-relative insufficiency from outright refusal-violation as two genuinely independent grounds for blocking a quotient. This chapter has closed the part by showing that even the targets of recovery themselves, not merely the tasks that call on them, form a structured lattice rather than a single scale, with history at the top, a well-behaved chain of successively coarser totals along one branch, and a pair of incomparable second-tier targets whose relationship depends on the substrate rather than on anything the general theory can settle in advance.

Part VI turns from this general apparatus to synthesis: a comparative survey across the substrates this book has not yet visited directly, ecology, infrastructure, organizations, law, machine learning, and language, followed by a treatment of how a claim built on this framework can actually be tested, and a closing statement of the book's two organizing principles in their most general form.

Exercises

1. Confirm each implication in Proposition 17.1 directly for a specific numerical example: choose values for U_1, U_2, R_{12}, compute the total, then show how ordering and sign recovery follow from a coarser summary than the total itself.

2. Construct a third example of extensional recovery without Ω-recovery, distinct from magnetic hysteresis, drawn from an economic or organizational setting.

3. In the structural-rule counterexample used in the proof of Proposition 17.2, describe explicitly what kind of admissibility rule would need to hold for Ω to depend only on structure and not on values, and give one concrete instance of such a rule within this book's Pop/Refuse/Bind/Collapse calculus.

4. A colleague claims that a particular data-reporting practice "preserves everything that matters." Using this chapter's lattice, list the specific questions you would need to ask before evaluating that claim, and explain why no single follow-up question suffices.

Part V

Cross-Substrate Synthesis

Chapter 18

A Survey Across Six Further Domains

18.1 What this chapter can and cannot do

Each chapter of Part IV derived a genuine result specific to its substrate: a proved identity (variance decomposition), a theorem borrowed and precisely specialized (Fisher–Neyman sufficiency), a cited external theorem given a rigorous role (the data processing inequality). This chapter covers six further domains in the space the original conception of this project had allotted seven full chapters, and it is honest to say plainly what that compression costs. Where Part IV proved things, this chapter mostly *identifies* things: for each domain, it names what plays the role of an addressable entry, a self-contribution U_i, and a relational contribution R_{ij}, and it gives exactly one fact, already established in that domain's own literature, precise and checkable, that the fact is a genuine instance of this book's binding-correction shape rather than a resemblance to it. What this chapter does not do, for any of the six domains, is rebuild that domain's admissibility theory, its refusal structure, or its recoverability lattice from scratch; those remain work for a future, longer treatment; here only the entry point is secured.

18.2 Ecology

Let the entries be species (or local populations), U_i the isolated growth rate of species i in the absence of all others, and R_{ij} the pairwise interaction coefficient between species i and j. The generalized Lotka–Volterra competition model,

$$\frac{dN_i}{dt} = r_i N_i \left(1 - \frac{N_i + \sum_{j \neq i} \alpha_{ij} N_j}{K_i}\right),$$

has an interaction matrix (α_{ij}) that is, precisely and without metaphor, a binding matrix in Chapter 11's sense: its sign at each pair records competition ($\alpha_{ij}, \alpha_{ji} > 0$), mutualism (both negative), or a predator–prey relationship (opposite signs, an asymmetric coupling of exactly the kind Chapter 11's directed-bindings section characterized as invisible to any symmetric quadratic evaluation). This is not an analogy constructed for this book; multispecies Lotka–Volterra models are standard theoretical ecology, and their interaction matrix is exactly the kind of object Chapters 4 and 11 already built the apparatus for.

18.3 Infrastructure

Let the entries be candidate facility sites, U_i a site's fixed or standalone cost, and R_{ij} the transportation or coordination cost that depends on the joint placement of sites i and j. The standard facility-location objective in operations research,

$$\sum_i f_i y_i + \sum_i \sum_j c_{ij} x_{ij},$$

separates exactly into a unary term (fixed costs f_i) and a pairwise term (flow costs c_{ij} depending on both a facility and a demand point, or on two facilities' relative placement), the same self-term/cross-term split this book has used since the Preface. A facility sited to minimize $\sum_i f_i y_i$ alone, ignoring the coupling terms, is exactly Chapter 15's profile-only unary collapse, and it is a genuinely common design error in practice, not merely a theoretical possibility.

18.4 Organizations

Let the entries be individual contributors on a project, U_i an individual's isolated output, and R_{ij} the coordination cost of every pair of contributors who must communicate. Brooks's observation, from software project management, that adding workers to a late project can make it later, rests on exactly the combinatorial fact established in Chapter 1: the number of pairwise communication channels among n people is $\binom{n}{2}$, growing quadratically in n, while total individual output capacity grows only linearly. This is precisely Chapter 1's diagonal/off-diagonal split, with the off-diagonal count $\binom{n}{2}$ doing the same work here that it did in the trinomial square, now attached to a cost rather than a value, and growing fast enough, as n increases, to overwhelm the linear gain the added workers were meant to provide.

18.5 Law

Law has already served as this book's running example for admissibility data specifically, since Chapter 8 introduced Refuse using two witnesses whose testimony traced to a common source. Nothing new needs to be built here; it is worth simply naming the correspondence explicitly now that the full apparatus is available. Entries are pieces of evidence, U_i the probative value of a piece considered alone, R_{ij} the corroboration or dependence between two pieces, and A, this book's admissibility component of the ledger, corresponds directly to rules of evidence: hearsay exclusions, privilege, and chain-of-custody requirements are all, in this book's vocabulary, recorded refusals, forbidding certain pieces of evidence from being identified with, or substituted for, certain others regardless of what any downstream verdict-computing task would otherwise tolerate. Chapter 16's two-fold admissibility theorem, task-relative sufficiency and refusal-respect as independent conditions, is close to a formal restatement of due process itself: a piece of evidence can be perfectly informative for the question at hand and still properly excluded, exactly Chapter 16's Example with condition (b) failing while condition (a) would have held.

18.6 Machine learning

Let the entries be positions in a sequence (tokens), U_i a token's representation considered alone, and R_{ij} the pairwise coupling a model computes between two positions. The self-attention mechanism at the center of the transformer architecture computes exactly such a matrix, QK^{T}, scoring every pair of positions before normalizing and combining; because Q (queries) and K (keys) are different learned projections of the same input, this matrix is not symmetric in general, $\left(QK^{\mathsf{T}}\right)_{ij} \neq \left(QK^{\mathsf{T}}\right)_{ji}$, making attention a directed binding matrix in exactly Chapter 11's sense, with the same consequence established there: a symmetric evaluation built on top of it would be blind to the antisymmetric part, and attention's asymmetry, token i attending to token j differently than j attends to i, is precisely the directional information such an evaluation would discard.

18.7 Language

Let the entries be words, U_i a word's context-free frequency, and R_{ij} the association between two words appearing together. *Pointwise mutual information*,

$$\mathrm{PMI}(w_i, w_j) = \log \frac{p(w_i, w_j)}{p(w_i)p(w_j)},$$

a standard quantity in computational linguistics for measuring how strongly two words co-occur beyond chance, is not merely analogous to Chapter 13's mutual information; it is its pointwise, unaveraged form, related exactly by $I(X;Y) = \sum_{x,y} p(x,y)\,\mathrm{PMI}(x,y)$, mutual information being precisely the expectation of pointwise mutual information under the joint distribution. Every qualification Chapter 13 attached to $I(X;Y)$ therefore applies to PMI as well, including the caution that a specific evaluation's silence, here a PMI near zero for a specific word pair, is evidence about that evaluation's sensitivity, not a general certificate that the two words are unrelated in every sense a linguist might care about.

18.8 A comparative summary

Domain	**Entries**	U_i	R_{ij}
Graphs (Ch. 11)	vertices	self-loop weight	edge weight a_{ij}
Probability (Ch. 12)	random variables	$\mathrm{Var}(X_i)$	$\mathrm{Cov}(X_i, X_j)$
Information (Ch. 13)	signals	$H(X_i)$	$I(X_i; X_j)$
Ecology	species	isolated growth rate	α_{ij} (Lotka–Volterra)
Infrastructure	facility sites	fixed cost f_i	flow cost c_{ij}
Organizations	contributors	individual output	coordination cost
Law	evidence items	probative value alone	corroboration/dependence
Machine learning	sequence positions	token representation	attention score $(QK^{\mathsf{T}})_{ij}$
Language	words	context-free frequency	$\mathrm{PMI}(w_i, w_j)$

Reading down this table domain by domain is not the point of building it. The point is that the same two columns, U_i and R_{ij}, recur across every row, and that in at least one precise, checkable respect per domain, the correspondence is not this book's invention but an existing,

independently developed fact about that domain, now visible as one more instance of the shape first derived from a square of side $x + y + z$.

18.9 What remains

None of the six sections above attempted to carry a domain's admissibility structure, its Refuse relation, its recoverability lattice, as far as Chapters 15 through 17 carried graphs, probability, and information. That would require, for each domain, the same kind of sustained, chapter-length attention Part IV gave its three substrates, work this condensed edition does not undertake. What this chapter has secured is narrower and, for a survey, sufficient: six further domains in which U_i and R_{ij} are not merely suggestive labels but names for quantities those domains' own specialists already compute, argue about, and build models from, which is exactly the condition Chapter 2 set at the very start of this book for taking an application seriously rather than merely gesturing at it.

Chapter 19 turns from identifying correspondences to testing them: given a proposed instance of this book's framework in some new domain, what would it take to design an experiment capable of showing the framework wrong.

Exercises

1. For the ecology example, identify a concrete three-species system (real or hypothetical) and classify each pairwise α_{ij} as competitive, mutualistic, or predatory, then determine whether the resulting interaction matrix is symmetric.

2. For the organizations example, compute $\binom{n}{2}$ for $n = 5, 10, 20$ and compare its growth rate to n itself, reconstructing the core of Brooks's argument numerically.

3. Using the identity $I(X;Y) = \sum_{x,y} p(x,y)\,\mathrm{PMI}(x,y)$, construct a small joint distribution over two binary variables for which some individual $\mathrm{PMI}(x,y)$ is negative even though $I(X;Y) > 0$ overall, and interpret what a negative PMI value means for that specific pair.

4. Choose one domain from this chapter and propose a plausible admissibility rule (a Refuse-type constraint) that domain's practitioners already observe informally, in the style of Chapter 16's due-process discussion of law.

Chapter 19

Designing a Flat Experiment

19.1 The additivity baseline

Every empirical claim this book's vocabulary can make, that two entries are bound, that a collapse is admissible, that a correction reflects a genuine relation, needs a stated null to be measured against, and this book has in fact been using one specific null since Chapter 2 without naming it as an experimental design principle. Call it the *additivity baseline*: the hypothesis that the separate evaluation $E_0(X) = \sum_i U_i$ correctly predicts the joint evaluation, equivalently that the binding correction vanishes. Any claim of interaction, coupling, or synergy is a claim about the residual

$$\epsilon_B = E(X) - E_0(X),$$

and it is worth stating plainly what the rest of this book has treated as background: a claim of "interaction" that is never checked against this specific baseline, computed and reported explicitly, is not yet a falsifiable claim in this book's sense. It is a description waiting to be turned into one.

19.2 Binding versus confounding

A nonzero ϵ_B, observed rather than merely posited, admits more explanations than Chapter 5's four-fold list considered, once the entries in question are allowed to have their own causal histories rather than being treated as bare algebraic symbols. A third variable Z, causing both x and y without any direct coupling between them, produces exactly the same observed correlation, and the same nonzero ϵ_B under a covariance-based Q, as a genuine direct binding would. This is confounding, and it is a distinct failure mode from any of Chapter 5's four, since it does not involve Q's sensitivity at all; Q may be perfectly well-chosen, and the correction it reports perfectly accurately measured, while still misattributing to a direct x–y relation what a hidden common cause actually produced.

Observation alone, however carefully measured, cannot in general distinguish a genuine binding from a confound. This is not a limitation of this book's specific apparatus; it is the central reason experimental design, as a discipline, exists at all, and it motivates the one methodological upgrade this chapter adds to everything built so far.

19.3 Intervention on relations

Chapter 5 left an exercise open, deliberately: given $\Delta_Q(x, y) = 0$ observed, and an intervention that changes x while holding the mechanism generating y fixed causing Δ_Q to become nonzero, which of Chapter 5's four explanations does this rule out, and which remain possible? The apparatus built since then makes a precise answer possible.

Proposition 19.1 (What an intervention on x establishes)**.** *Let $\Delta_Q(x, y) = 2B(x, y)$ for a fixed bilinear B, suppose $\Delta_Q(x_0, y) = 0$ is observed at some value x_0, and suppose an intervention changes x to $x_1 \neq x_0$, holding y and the coupling B itself fixed, after which $\Delta_Q(x_1, y) \neq 0$ is observed. Then:*

(a) B is not identically zero: some pair of values makes B nonzero, so "genuine independence," read as $B \equiv 0$, is ruled out.

(b) The original vanishing at x_0 was a local fact, in the exact sense of the remark following Proposition 2.3, not evidence that Q is insensitive to this coupling in general, since the same Q now registers a nonzero correction.

(c) The experiment does not rule out that further coupling mechanisms, beyond B, exist and happened to cancel B's contribution exactly at x_0, if the evaluation in use is a sum of several such mechanisms rather than a single bilinear form.

(d) The experiment does not rule out that a smaller, genuine dependence existed at x_0 that fell below the resolution of whatever instrument produced the original measurement, since establishing that requires independent evidence about the instrument, not supplied by this experiment alone.

Proof. (a): if $B \equiv 0$, then $\Delta_Q(x, y) = 2B(x, y) = 0$ for every x, so no change in x alone could produce a nonzero result; the observation directly contradicts this. (b): the same Q, hence the same sensitivity to the type of coupling present, is in use before and after the intervention, and it registers a nonzero value after; the earlier zero is therefore a fact about the specific point x_0 (Proposition 2.3's remark), not about Q's general capacity to detect this coupling. (c) and (d) are non-results: neither is contradicted by the observation, since a cancellation specific to x_0 or a below-threshold dependence at x_0 are both fully compatible with a larger, detectable Δ_Q once x moves away from x_0. □

This is a sharper resolution than a first pass at the exercise might suggest. It is tempting to read a successful intervention as ruling out all four of Chapter 5's explanations at once, since something was clearly detected; Proposition 19.1 shows this overstates the case. Genuine independence and a globally undetecting Q are cleanly ruled out, because both would have made the intervention's result impossible in principle. Cancellation and coarseness are not ruled out at all, because both remain fully consistent with exactly the result observed. An intervention is a genuinely stronger form of evidence than a passive correlation, but it is stronger in a specific, limited way, ruling out two of four explanations rigorously rather than replacing observation with certainty across the board, and reporting an intervention's result as having "proven the relationship is real" without this qualification would overstate what Proposition 19.1 actually delivers.

19.4 Collapse audits

Chapters 15 through 17 gave admissibility and recoverability a formal vocabulary; this section turns that vocabulary into a checklist. Given a proposed collapse T_S and a set of candidate downstream tasks $q_1, \ldots, q_m$, a *collapse audit* checks, for each q_k, whether T_S is q_k-sufficient (Theorem 15.2) and whether T_S respects every recorded refusal (Theorem 16.2, condition (b)). The output of an audit is not a single verdict but a table: which tasks the collapse safely serves, which it does not, and which refusals, if any, it would violate regardless of task. Chapter 16's worked three-step procedure is the audit template for quotients specifically; nothing in this section adds new theory, only the discipline of running the existing checks systematically, against a stated list of tasks, rather than against whichever task happens to be uppermost in mind when the collapse is first proposed.

19.5 The flat experiment

The pieces above combine into a template, deliberately called *flat* because it strips a proposed claim of binding down to the smallest structure capable of testing it, no more elaborate than the claim requires.

> A flat experiment specifies: (1) the **components**, the entries x, y (or more) whose relation is in question; (2) the **baseline**, the additivity prediction $E_0 = \sum_i U_i$ against which any residual will be measured; (3) the **binding rule**, a precise, ideally interventable specification of how the claimed coupling is supposed to operate; (4) the **evaluation**, the specific Q or σ used to measure the outcome, named in advance; and (5) the **falsification condition**, a statement of what result would count as evidence against the claimed binding, fixed before the experiment is run.

The fifth element is the one most often missing from applications of this book's vocabulary offered informally, and its absence is what turns a real claim into an unfalsifiable one. "x and y interact" is not yet a flat experiment; "if the intervention described in step 3 fails to produce a residual exceeding the noise floor of the evaluation in step 4, the claimed binding is rejected" is. Chapter 18's six domains each gestured toward a claim of this shape, an interaction matrix, a coordination cost, an attention score, without specifying the falsification condition that would make any one of them a tested claim rather than a suggestive correspondence; supplying that condition, for any specific instance a reader wants to pursue beyond this book, is exactly the missing step this chapter has tried to make explicit.

19.6 Closing remarks

This chapter has not introduced new mathematics; every proposition in it specializes machinery built in earlier chapters. What it has added is discipline: a named baseline against which any claim of interaction must be measured, a precise account of what intervention does and does not establish (finally resolving the question Chapter 5 left open, and resolving it more narrowly than a first guess would suggest), a systematic procedure for auditing a proposed collapse against a stated list of tasks, and a template requiring every claim built on this book's vocabulary to state, in advance, what would count as its own refutation. Chapter 20 closes the book by stating, one

final time and in their most general form, the two principles every chapter since the Preface has been assembling evidence for.

Exercises

1. Design a flat experiment, using the five-part template of this chapter, for one of Chapter 18's six domain correspondences (your choice), specifying all five elements explicitly.

2. Construct a scenario in which a genuine confound (a variable Z causing both x and y) produces exactly the same measured $\Delta_Q(x, y)$ as a direct binding would, and describe an intervention that would distinguish the two cases.

3. Using Proposition 19.1, explain precisely why an intervention that successfully produces a nonzero Δ_Q after an initial zero reading does *not*, by itself, rule out that measurement coarseness was part of the original explanation.

4. Perform a collapse audit, in the sense of this chapter, on the profile-only unary collapse from Chapter 15's taxonomy, against three tasks of your choosing, at least one of which the collapse should fail.

Chapter 20

Distinction and Continuation

20.1 Return to the rectangle

The Preface opened with a square of side $x + y + z$, divided into three colored squares and six uncolored rectangles, and asked where the rectangles come from. Nineteen chapters later, the honest answer is no longer the geometric one. The rectangles are not a fact about area. They are the simplest visible instance of a fact this book has now established, specialized, and cross-checked across a dozen substrates: that evaluating a joint structure and summing separate evaluations of its parts are, in general, two different computations, that the discrepancy between them is a specific, well-defined, and often recoverable quantity, and that what a given representation retains or discards of that quantity is a decision, not a byproduct, with consequences for exactly which future questions the representation can still answer.

Three principles, each stated informally at some point along the way, can now be stated in the general form the book has actually earned.

20.2 The binding principle

> **The Binding Principle.** Whenever an evaluation is non-additive over a decomposition into distinguishable components, the discrepancy between the joint evaluation and the sum of separate evaluations is a specific, computable quantity, not a residual to be ignored or a default sign of a mysterious surplus. For any evaluation E and components $x_1, \ldots, x_n$,
>
> $$E\left(\bigoplus_i x_i\right) = \sum_i E(x_i) + \Delta_E(x_1, \ldots, x_n),$$
>
> and the burden of any substantive claim built on this decomposition is to characterize Δ_E, not merely to observe that it is nonzero.

This is Chapter 2's definition, restated with the qualifications the intervening chapters earned it. Δ_E can be positive, negative, or zero (Chapter 2); when E is quadratic, it is exactly twice a bilinear form, recoverable from evaluations of E alone (Chapters 4 and 5); it can be diagonalized away in one basis while remaining fully recoverable from another, provided the change of coordinates is retained (Chapter 6, corrected); a nonzero or zero Δ_E is never, by itself, a

certificate of the presence or absence of a specific underlying mechanism, only a fact about E at that specific point (Chapter 5, and Chapter 19's Proposition 19.1 sharpening exactly how much an intervention can and cannot establish about it); and it names the same shape of quantity in quadratic forms, graphs, variance, entropy, and, more provisionally, six further domains this book surveyed but did not fully develop (Chapters 11 through 18). The principle is not that binding produces surplus. It is that binding produces a term, of a specific and analyzable shape, whose content has to be derived, measured, or intervened upon, never assumed.

20.3 The provenance principle

> **The Provenance Principle.** Equal totals do not imply equal derivations. For any evaluation σ and any two generative histories H_1, H_2,
>
> $$\sigma(H_1) = \sigma(H_2) \implies H_1 = H_2,$$
>
> and a representation retaining only $\sigma(H)$ cannot, in general, answer any question that depends on which H actually produced it.

This is Chapter 3's Proposition 3.2, and its consequences occupied a large fraction of everything that followed. It is why the sample mean, sufficient for the mean of a Gaussian, discards which specific sample produced it (Chapter 12); why a graph quotient's aggregate total cannot answer a question about one specific edge inside it (Chapter 11); why a corrected entry cannot, in general, be patched into a collapsed scalar without reconstructing the relation it was bound through (Chapter 14); and why the interaction ledger's fourth component, H itself, sits at the unique top of the recoverability lattice, strictly above everything a scalar total, an ordering, or even a full extensional profile can offer on its own (Chapter 17). The principle is not a counsel of despair about summarization; most tasks want exactly a summary and nothing more. It is a discipline about knowing, in advance, which of one's tasks are among those most, and which are not.

20.4 The admissible-collapse principle

> **The Admissible-Collapse Principle.** A collapse is justified only relative to a stated downstream requirement. For a transformation T and a task q with observation map r_q,
>
> $$T \text{ is admissible for } q \iff T(H_1) = T(H_2) \implies r_q(\omega(H_1)) = r_q(\omega(H_2)),$$
>
> and no collapse is admissible or inadmissible *simpliciter*, only relative to some named q, some named refusal set A, or both.

This is Theorem 10.2, by way of Corollary 8.6 and Chapter 16's two-fold theorem separating task-relative insufficiency from outright refusal-violation as independent grounds for blocking a collapse. It is the principle that makes "preserves everything that matters" an unanswerable question until *matters for what* has been supplied, that makes non-injectivity a candidate for information loss rather than a synonym for it (Chapter 10's hierarchy, Chapter 11's standing convention), and that gives due process, audit, peer review, and version control the same formal shape: institutions, in this book's vocabulary, are largely machinery for making q explicit before a collapse is permitted, rather than trusting a collapse chosen for convenience to happen to serve whatever q turns out to matter later.

20.5 Why continuation

The three principles share a single orientation, worth naming once more before the book closes. None of them is ultimately about getting the present computation right, though each has quite a lot to say about that. Each is about what a representation permits *next*: whether a future audit can find what it needs (provenance), whether a future evaluation can be interpreted correctly (binding), and whether a future task will discover that an earlier convenience quietly foreclosed it (admissible collapse). This is why Chapter 7 introduced worlds as pairs (H, Ω) rather than simply as states, and why the recoverability lattice of Chapter 17 has Ω as one of its two second-tier targets, incomparable with the extensional profile rather than subordinate to it: a system's history matters not only for what it currently is, but for what it can still become, and a representation is judged, throughout this book, by whether it keeps that question answerable, not merely by whether it gets today's number right.

20.6 What this book has and has not shown

It is worth closing with the same discipline the book has tried to apply to every claim inside it. Parts I through III proved general, substrate-independent theorems: the binding correction's basic algebra, its bilinear specialization, the formal apparatus of worlds and the four operators, and the general sufficiency criterion unifying all of it. Part IV specialized that apparatus rigorously to three substrates, graphs, probability, and information theory, each with genuine proofs, a real correspondence to independently developed theory in the probability case, and honest disclosure of exactly which results were derived internally and which were cited as standard external theorems. Part V generalized the admissibility question itself, giving it the interaction ledger, a genuine two-fold theorem for quotienting, and a proven, non-trivial lattice structure for recoverability. Part VI is the part to be most careful about overstating: its survey of six further domains identified real, checkable correspondences, one flagship fact per domain, but did not carry any of them as far as Part IV carried its three, and said so explicitly; its treatment of falsifiability supplied a template and one substantial worked resolution, not a general theory of experimental design.

This book has not shown that admissibility, binding, and provenance are the only concepts worth this kind of attention, nor that every domain in Chapter 18's table would sustain the same weight of theory that graphs, probability, and information theory did. What it has shown, with the rigor each chapter's corrections and counter-checks were meant to earn rather than assume, is that a single algebraic fact, visible already in a square divided into three colors and six rectangles, generalizes further, and more usefully, than its elementary origin would suggest: that self-terms and cross-terms, provenance and collapse, and the question of what a representation still permits, are the same handful of distinctions, recurring under new names, in every structured system this book has had the space to examine closely enough to check.