

## *How Analogy Makes Meaning*

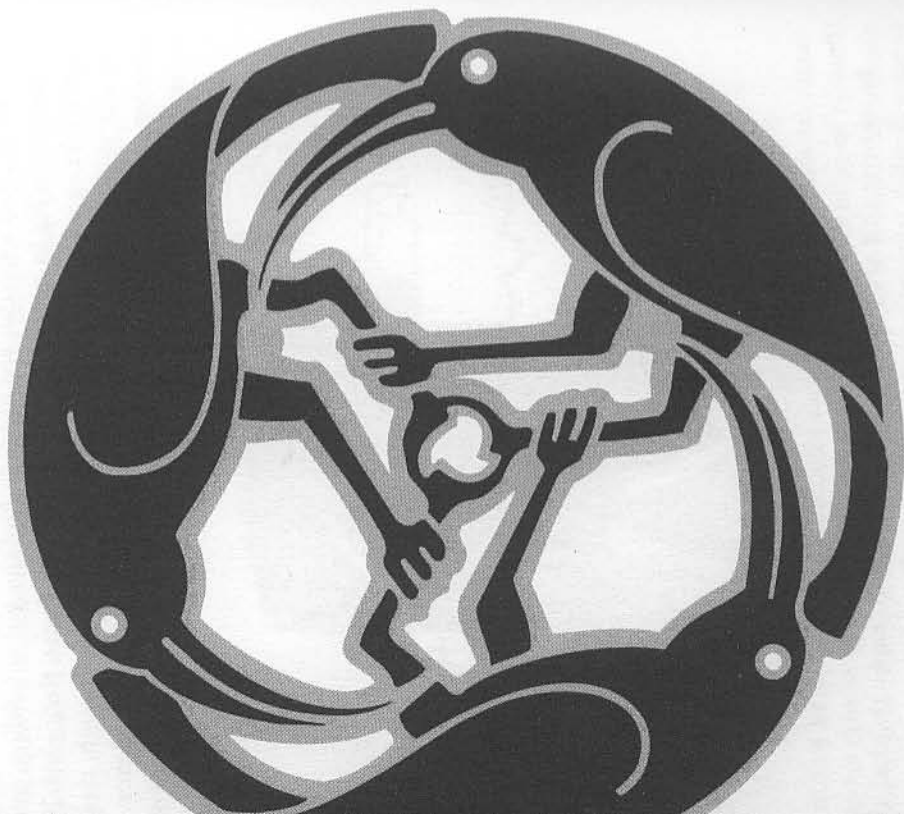


### The Double Aboutness of Formulas in PM

**IMAGINE** the bewilderment of newly knighted Lord Russell when a young Austrian Turk named “Kurt” declared in print that *Principia Mathematica*, that formidable intellectual fortress so painstakingly erected as a bastion against the horrid scourge of self-referentiality, was in fact riddled through and through with formulas allegedly stating all sorts of absurd and incomprehensible things about themselves. How could such an outrage ever have been allowed to take place? How could vacuously twitting self-referential propositions have managed to sneak through the thick ramparts of the beautiful and timeless Theory of Ramified Types? This upstart Austrian sorcerer had surely cast some sort of evil spell, but by what means had he wrought his wretched deed?

The answer is that in his classic article — “On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems (I)” — Gödel had re-analyzed the notion of *meaning* and had concluded that what a formula of *PM* meant was not so simple — not so unambiguous — as Russell had thought. To be fair, Russell himself had always insisted that *PM*’s strange-looking long formulas had *no* intrinsic meaning. Indeed, since the theorems of *PM* were churned out by formal rules that paid no attention to meaning, Russell often said the whole work was just an array of meaningless marks (and as you saw at the end of Chapter 9, the pages of *Principia Mathematica* often look more like some exotic artwork than like a work of math).

And yet Russell was also careful to point out that all these curious patterns of horseshoes, hooks, stars, and squiggles *could* be interpreted, if one wished, as being statements about numbers and their properties,



because under duress, one could read the meaningless vertical egg '0' as standing for the number zero, the equally meaningless cross '+' as standing for addition, and so on, in which case all the theorems of *PM* came out as statements about numbers — but not just random blatherings about them. Just imagine how crushed Russell would have been if the squiggle pattern "ss0 + ss0 = sssss0" turned out to be a theorem of *PM*! To him, this would have been a disaster of the highest order. Thus he had to concede that there *was* meaning to be found in his murky-looking tomes (otherwise, why would he have spent long years of his life writing them, and why would he care which strings were theorems?) — but that meaning depended on using a *mapping* that linked shapes on paper to abstract magnitudes (e.g., zero, one, two...), operations (e.g., addition), relationships (e.g., equality), concepts of logic (e.g., "not", "and", "there exists", "all"), and so forth.

Russell's dependence on a systematic mapping to read meanings into his fortress of symbols is quite telling, because what the young Turk Gödel had discovered was simply a *different* systematic mapping (a much more complicated one, admittedly) by which one could read *different* meanings into the selfsame fortress. Ironically, then, Gödel's discovery was very much in the Russellian spirit.

By virtue of Gödel's subtle new code, which systematically mapped strings of symbols onto numbers and vice versa (recall also that it mapped typographical shunting laws onto numerical calculations, and vice versa), many formulas could be read on a second level. The first level of meaning, obtained via the old standard mapping, was always about numbers, just as Russell claimed, but the second level of meaning, using Gödel's newly revealed mapping (piggybacked on top of Russell's first mapping), was about *formulas*, and since both levels of meaning depended on mappings, Gödel's new level of meaning was no less real and no less valid than Russell's original one — just somewhat harder to see.

### Extra Meanings Come for Free, Thanks to You, Analogy!

In my many years of reflecting about what Gödel did in 1931, it is this insight of his into the roots of meaning — his discovery that, thanks to a mapping, full-fledged meaning can suddenly appear in a spot where it was entirely unsuspected — that has always struck me the most. I find this insight as profound as it is simple. Strangely, though, I have seldom if ever seen this idea talked about in a way that brings out the profundity I find in it, and so I've decided to try to tackle that challenge myself in this chapter. To this end, I will use a series of examples that start rather trivially and grow in subtlety, and hopefully in humor as well. So here we go.

Standing in line with a friend in a café, I spot a large chocolate cake on a platter behind the counter, and I ask the server to give me a piece of it. My friend is tempted but doesn't take one. We go to our table and after my first bite of cake, I say, "Oh, this tastes awful." I mean, of course, not merely that my one slice is bad but that the whole cake is bad, so that my friend should feel wise (or lucky) to have refrained. This kind of mundane remark exemplifies how we effortlessly generalize outwards. We unconsciously think, "This piece of the cake is very much like the rest of the cake, so a statement about it will apply equally well to any other piece." (There is also another analogy presumed here, which is that my friend's reaction to foods is similar to mine, but I'll leave that alone.)

Let's try another example, just a tiny bit more daring. There's a batch of cookies on a plate at a party and I pick one up, take a bite, and remark to my children, "This is delicious!" Immediately, my kids take one each. Why? Because they wanted to taste something delicious. Yes, but how did they jump from my statement about *my* cookie to a conclusion about *other* cookies on the plate? The obvious answer is that the cookies are all "the same" in some sense. Unlike the pieces of cake, though, the cookies are not all parts of one single physical object, and thus they are ever so slightly "more different" from one another than are the pieces of cake — but they were made by the same person from the same ingredients using the same equipment. These cookies come from a single batch — they belong to the same category. In all relevant aspects, we see them as interchangeable. To be sure, each one is unique, but in the senses that count for human cookie consumption, they are almost certain to be equivalent. Therefore if I say about a particular one, "My, this is delicious!", my statement's meaning implicitly jumps across to any other of them, by the force of analogy. Now, to be sure, it's a rather trivial analogy to jump from one cookie to another when they all come from the same plate, but it's nonetheless an analogy, and it allows my specific statement "This is delicious!" to be taken as a general statement about all the cookies at once.

You may find these examples too childish for words. The first one involves an "analogy" between several slices of the same cake, and the second one an "analogy" between several cookies on the same plate. Are these banalities even worthy of the label "analogy"? To me there is no doubt about it; indeed, it is out of a dense fabric of a myriad of invisible, throwaway analogies no grander than these that the vast majority of our rich mental life is built. Yet we take such throwaway analogies so much for granted that we tend to think that the word "analogy" must denote something far more exalted. But one of my life's most recurrent theme

songs is that we should have great respect for what seem like the most mundane of analogies, for when they are examined, they often can be seen to have sprung from, and to reveal, the deepest roots of human cognition.

### Exploiting the Analogies in Everyday Situations

As we've just seen, a remark made with the aim of talking about situation A can also implicitly apply to situation B, even if there was no intention of talking about B, and B was never mentioned at all. All it takes is that there be an easy analogy — an unforced mapping that reveals both situations to have essentially the same central structure or conceptual core — and then the extra meaning is there to be read, whether one chooses to read it or not. In short, a statement about one situation can be heard as if it were about an analogous — or, to use a slightly technical term, *isomorphic* — situation. An isomorphism is just a formalized and strict analogy — one in which the network of parallelisms between two situations has been spelled out explicitly and precisely — and I'll use the term freely below.

When an analogy between situations A and B is glaringly obvious (no matter how simple it is), we sometimes will exploit it to talk “accidentally on purpose” about situation B by pretending to be talking only about situation A. “Hey there, Andy — take your muddy boots off when you come into the house!” Such a sentence, when shouted at one's five-year-old son who is tramping in the front door with his equally mud-oozing friend Bill, is obviously addressed just as much to Bill as to Andy, via a very simple, very apparent analogy (a boy-to-boy leap, if you will, much like the earlier cookie-to-cookie leap). Hinting by analogy allows us to get our message across politely but effectively. Of course we have to be pretty sure that the person at whom we're beaming our implicit message (Bill, here) is likely to be aware of the A/B analogy, for otherwise our clever and diplomatic ploy will all have been for naught.

Onward and upward in our chain of examples. People in romantic situations make use of such devices all the time. One evening, at a passionate moment during a tender clinch, Xerxes queries of his sweetie pie Yolanda, “Do I have bad breath?” He genuinely wants to know the answer, which is quite thoughtful of him, but at the same time his question is loaded (whether he intends it to be or not) with a second level of meaning, one not quite so thoughtful: “You have bad breath!” Yolanda answers his question but of course she also picks up on its potential alternative meaning in a flash. In fact, she suspects that Xerxes' *real* intent was to tell her about *her* breath, not to find out about his own — he was just being diplomatic.

Now how can one statement speak on two levels at once? How can a second meaning lie lurking inside a first meaning? You know the answer as well as I do, dear reader, but let me spell it out anyway. Just as in the muddy-boots situation, there is a very simple, very loud, very salient, very obvious analogy between the two parties, and this means that any statement made about X will be (or at least can be) heard as being about Y at the same time. The X/Y mapping, the analogy, the partial isomorphism — whatever you wish to call it — carries the meaning efficiently and reliably from one framework over to the other.

Let's look at this mode of communication in a slightly more delicate romantic situation. Audrey, who is not sure how serious Ben is about her, “innocently” turns the conversation to their mutual friends Cynthia and Dave, and “innocently” asks Ben what he thinks of Dave's inability to commit to Cynthia. Ben, no fool, swiftly senses the danger here, and so at first he is wary about saying anything specific since he may incriminate himself even though talking “only” about Dave, but then he also realizes that this danger gives him an opportunity to convey to Audrey some things that he hasn't dared to raise with her directly. Accordingly, Ben replies with a calculated air of nonchalance that he can imagine why Dave might be hesitant to commit himself, since, after all, Cynthia is so much more intellectual than Dave is. Ben is hoping that Audrey will pick up on the hint that since *she* is so much more involved in art than *he* is, that's why *he's* been hesitant to commit himself as well. His hint is carried to her implicitly but clearly via the rather strong couple-to-couple analogy that both Audrey and Ben have built up in their heads over the past several months without ever breathing a word of it to each other. Ben has managed to talk very clearly about himself although without ever talking *directly* about himself, and what's more, both he and Audrey know this is so.

The preceding situation might strike you as being very contrived, thus leaving you with the impression that seeing one romantic situation as “coding” for another is a fragile and unlikely possibility. But nothing could be further from the truth. If two people are romantically involved (or even if they aren't, but at least one of them feels there's a potential spark), then almost any conversation between them about any romance whatsoever, no matter who it involves, stands a good chance of being heard by one or both of them as putting a spotlight on their own situation. Such boomeranging-back is almost inevitable because romances, even very good ones, are filled with uncertainty and yearning. We are always on the lookout for clues or insights into our romantic lives, and analogies are among the greatest sources of clues and insights. Therefore, to notice an analogy between

ourselves and another couple that is occupying center stage in our conversation is pretty much a piece of cake handed to us on a silver platter. The crucial question is whether it tastes good or not.

### The Latent Ambiguity of the Village Baker's Remarks

Indirect reference of the sort just discussed is often artistically exploited in literature, where, because of a strong analogy that readers easily perceive between Situations A and B, lines uttered by characters in Situation A can easily be heard as applying equally well to Situation B. Sometimes the characters in Situation A are completely unaware of Situation B, which can make for a humorous effect, whereas other times the characters in Situation A are simultaneously characters in Situation B, but aren't aware of (or aren't thinking about) the analogy linking the two situations they are in. The latter creates a great sense of irony, of course.

Since I recently saw a lovely example of this, I can't resist telling you about it. It happens at the end of the 1938 film by Marcel Pagnol, *La Femme du boulanger*. Towards his wife Aurélie, who ran off with a local shepherd only to slink guiltily home three days later, Aimable, the drolly-named village baker, is all sweetness and light — but toward his cat Pomponnette, who, as it happens, *also* ran off and abandoned her mate Pompon three days earlier and who *also* came back on the same day as did Aurélie (all of this happening totally by coincidence, of course), Aimable is absolutely merciless. Taking the side of the injured Pompon (some might say “identifying with him”), Aimable rips Pomponnette to shreds with his accusatory words, and all of this happens right in front of the just-returned Aurélie, using excoriating phrases that viewers might well have expected him to use towards Aurélie. As if this were not enough, Aurélie consumes the heart-shaped bread that Aimable had prepared for himself for dinner (he had no inkling that she would return), while at the very same time, Pomponnette the straying kitycat, wearing a collar with a huge heart on it, is consuming the food just laid out for her mate Pompon.

Does Aimable the baker actually perceive the screamingly obvious analogy? Or could he be so kind and forgiving a soul that he doesn't see Aurélie and Pomponnette as two peas in a pod, and could the deliciously double-edged bile that we hear him savagely (but justifiably) dumping on the cat be innocently single-edged to him?

Whichever may be the case, I urge you to go out and see the film; it's a poignant masterpiece. And if by some strange chance your very own sweetheart, sitting at your side and savoring the movie with you, has just returned to the nest after *une toute petite amourlette* with some third party, just

imagine how she or he is going to start squirming when that last scene arrives! But why on earth would someone *outside* the movie feel the sting of a volley of stern rebukes made by someone *inside* the movie? Ah, well... analogy has force in proportion to its precision and its visibility.

### Chantal and the Piggybacked Levels of Meaning

Let's now explore an analogy whose two sides are more different than two cookies or two lovers, more different even than a straying wife and a straying cat. It's an analogy that comes up, albeit implicitly, when we are watching a video on our TV — let's say, a show about a French baker, his wife, his friends, and his cats. The point is that we are not *really* watching the cavortings of people and cats — not literally, anyway. To say that we are doing so is a useful shorthand, since what we are actually seeing is a myriad of pixels that are copying, in a perfect lockstep-synchronized fashion, dynamically shifting patterns of color splotches that once were scattered off some animate and inanimate objects in a long-ago-and-far-away French village. We are watching a million or so dots that “code” for those people's actions, but luckily the code is very easy for us to decode — so totally effortless, in fact, that we are sucked in by the mapping, by the isomorphism (the screen/scene *analogy*, if you will), and we find ourselves “teleported” to some remote place and time where we seem to be seeing events happening just as they normally do; we feel it is annoyingly nitpicky to make fine distinctions about whether we are “really” watching those events or not. (Are we *really* talking to each other if we talk by phone?)

It is all too easy to forget that moths, flies, dogs, cats, neonates, television cameras, and other small-souled beings do not perceive a television screen as we do. Although it's hard for us to imagine, they see the pixels in a raw, uninterpreted fashion, and thus to them a TV screen is as drained of long-ago-and-far-away meanings as is, to you or me, a pile of fall leaves, a Jackson Pollock painting, or a newspaper article in Malagasy (my apologies to you if you speak Malagasy; in that case, please replace it by Icelandic — and don't tell me that you speak that language, too!). “Reading” a TV screen at the representational level is intellectually far beyond such creatures, even if for most humans it is essentially second nature already by age two or so.

A dog gazing vacantly at a television screen, unable to make out any imagery, unaware even that any imagery is intended, is thus not unlike Lord Russell staring blankly at a formula of his beloved system *PM* and seeing only its “easy” (arithmetical) meaning, while the other meaning, the mapping-mediated meaning due to Gödel, lies intellectually beyond him,

utterly inaccessible, utterly undreamt-of. Or perhaps you think this comparison is unfair to Sir Bertrand, and in a way I agree, so let me try to make it a little more realistic and more generous.

Instead of a dog that, when placed in front of a TV screen, sees only pixels rather than people, imagine little three-year-old Chantal Duplessix, who is watching *La Femme du boulanger* with her parents. All three are native speakers of French, so there's no language barrier. Just like her *maman* and *papa*, Chantal sees right through the pixels to the events in the village, and when that wonderful final scene arrives and Aimable rakes the cat over the coals, Chantal laughs and laughs at Aimable's fury — but she doesn't suspect for a moment that there is *another* reading of his words. She's too young to get the analogy between Aurélie and Pomponette, and so for her there is only one meaning there. Filmmaker Pagnol's analogy-mediated meaning, which takes for granted the “simple” (although dog-eluding) mapping of pixels to remote events and thus piggybacks on it, is effortlessly perceived by her parents, but for the time being, it lies intellectually beyond Chantal, and is utterly inaccessible to her. In a few years, of course, things will be different — Chantal will have learned how to pick up on analogies between all sorts of complex situations — but that's how things are now.

With this situation, we can make a more realistic and more generous comparison to Bertrand Russell (yet another analogy!). Chantal, unlike a dog, does not merely see meaningless patterns of light on the screen; she effortlessly sees people and events — the “easy” meaning of the patterns. But there is a second level of meaning that takes the people and events for granted, a meaning transmitted by an analogy between events, and it's that *higher* level of meaning that eludes Chantal. In much the same way, Gödel's higher level of meaning, mediated by his mapping, his marvelous analogy, eluded Bertrand Russell. From what I have read about Russell, he never saw the second level of meaning of formulas of *PM*. In a certain sad sense, the good Lord never learned to read his own holy books.

### Pickets at the Posh Shop

As I suggested above, your recently returned roving sweetheart might well hear an extra level of meaning while listening to Aimable chasise Pomponette. Thus a play or film can carry levels of meaning that the author never dreamt of. Let's consider, for example, the little-known 1931 play *The Posh Shop Picketers*, written by social activist playwright Rosalyn Wadhead (ever hear of her?). This play is about a wildcat strike called by the workers at Alf and Bertie's Posh Shop (I admit, I never did figure out what they sold there). In this play, there is a scene where shoppers

approaching the store's entrance are exhorted not to cross the picket line and not to buy anything in the store (“Alf and Bertie are filthy dirty! Please don't cross our Posh Shop pickets! Please cross over to the mom-and-pop shop!”). In the skilled hands of our playwright, this simple situation led to a drama of great tension. But for some reason, just before the play was to open, the ushers in the theater and the actors in the play got embroiled in a bitter dispute, as a result of which the ushers' union staged a wildcat strike on opening night, put up picket lines, and beseeched potential playgoers not to cross their lines to see *The Posh Shop Picketers*.

Obviously, given this unanticipated political context, the lines uttered by the actors inside the play assumed a powerful second meaning for viewers in the audience, an extra level of meaning that Rosalyn Wadhead never intended. In fact, the picketing Posh Shop worker named “Cagey”, who disgustedly proclaims, after a brash matron pushes her aside and arrogantly strides into Alf and Bertie's upscale showroom, “Anyone who crosses the picket line in front of Alf and Bertie's Posh Shop is scum”, was inevitably heard by everyone in the audience (which by definition consisted solely of people who had crossed the picket line outside the theater) as saying, “Anyone who crossed the picket line in front of this theater is scum”, and of course this amounted to saying, “Anyone who is now sitting in this audience is scum”, which could also be heard as “You should not be listening to these lines”, which was the diametric opposite of what all the actors, including the one playing the part of Cagey, wanted to tell their audience, whose entry into the theater they so much appreciated, given the ushers' hostile picket line.

But what could the actors do about the fact that they were unmistakably calling their deeply appreciated audience “scum” and insinuating that no one should even have been there to hear these lines? Nothing. They *had* to recite the play's lines, and the analogy was there, it was blatant and strong, and therefore the ironic, twisting-back, self-referential meaning of Cagey's line, as well as of many others in the play, was unavoidable. Admittedly, the self-reference was *indirect* — mediated by an analogy — but that did not make it any less real or strong than would “direct” reference. Indeed, what we might be tempted to call “direct” reference is mediated by a code, too — the code between words and things given to us by our native language (Malagasy, Icelandic, etc.). It's just that *that* code is a simpler one (or at least a more familiar one). In sum, the seemingly sharp distinction between “direct” reference and “indirect” reference is only a matter of degree, not a black-and-white distinction. To repeat, analogy has force in proportion to its precision and visibility.

### Prince Hyppia: Math Dramatica

Well, so much for Rosalyn Wadhead and the surprise double-edgedness of the lines in *The Post Shop Picketers*, admittedly a rather obscure work. Let's move on to something completely different. We'll talk instead about the world-famous play *Prince Hyppia: Math Dramatica*, penned in the years 1910–1913 by the celebrated British playwright Y. Ted Enrustle (surely you've heard of him!). Fed up with all the too-clevah-by-hahf plays-about-plays that were all the rage in those days, he set out to write a play that would have nothing whatsoever to do with playwrighting or acting or the stage. And thus, in this renowned piece, as you doubtless recall, all the characters are strictly limited to speaking about various properties, from very simple to quite arcane, of whole numbers. How could anyone possibly get any further from writing a play about a play? For example, early on in Act I, the beautiful Princess Bloppia famously exclaims, "7 times 11 times 13 equals 1001!", to which the handsome Prince Hyppia excitedly retorts, "Wherefore the number 1001 is composite and not prime!" Theirs would seem to be a math made in heaven. (You may now groan.)

But it's in Act III that things really heat up. The climax comes when Princess Bloppia mentions an arithmetical fact about a certain very large integer  $g$ , and Prince Hyppia replies, "Wherefore the number  $g$  is saucy and not prim!" (It's a rare audience that fails to gasp in unison when they hear Hyppia's most math-dramatical outburst.) The curious thing is that the proud Prince seems to have no idea of the import of what he is saying, and even more ironically, apparently the playwright, Y. Ted Enrustle, didn't either. However, as everyone today knows, this remark of Prince Hyppia asserts — via the intermediary link of a tight analogy — that a certain long line of typographical symbols is "unpennable" using a standard set of conventions of dramaturgy that held, way back in those bygone days. And the funny thing is that the allegedly unpennable line is none other than the proclamation that the actor playing Prince Hyppia has just pronounced!

As you can well imagine, although Y. Ted Enrustle was constantly penning long lines of symbols that adhered to popular dramaturgical conventions (after all, that was his livelihood!), he'd never dreamt of a connection between the natural numbers (whose peculiar properties his curious characters accurately articulated) and the humble lines of symbols that he penned for his actors to read and memorize. Nonetheless, when, nearly two decades later, this droll coincidence was revealed to the play-going public in a wickedly witty review entitled "On Formerly Unpennable Proclamations in *Prince Hyppia: Math Dramatica* and Related Stageplays (I)"; authored by the acerbic young Turko-Viennese drama critic Gerd Kùlot

(I'll skip the details here, as the story is so well known), its piercing cogency was immediately appreciated by many, and as a result, playgoers who had read Kùlot's irreverent review became able to rehear many of the famous lines uttered in *Prince Hyppia: Math Dramatica* as if they were not about numbers at all, despite what Y. Ted Enrustle had intended, but were direct (and often quite biting) comments about Y. Ted Enrustle's play itself!

And thus it wasn't long before savvy audiences were reinterpreting the droll remarks by the oddball numerologist Qéc Dzhii (a character in *Prince Hyppia: Math Dramatica* who had gained notoriety for her nearly nonstop jabbering about why she preferred saucy numbers to prim numbers) as revealing, via allusions that now seemed hilariously obvious, why she preferred dramatic lines that were unpennable (using the dramaturgical conventions of the day) to lines that were pennable. Drama lovers considered this new way of understanding the play too delicious for words, for it revealed *Prince Hyppia* to be a play-about-a-play (with a vengeance!), although most of the credit for this insight was given to the brash young foreign critic rather than to the venerable elder playwright.

Y. Ted Enrustle, poor fellow, was simply gobsmacked — there's no other word for it. How could anyone in their right mind take Qéc Dzhii's lines in this preposterous fashion? They were only about numbers! After all, to write a drama that was about numbers and *only* about numbers had been his sole ambition, and he had slaved away for years to accomplish that noble goal!

Y. Ted Enrustle lashed out vehemently in print, maintaining that his play was decidedly *not* about a play, let alone about itself! Indeed, he went so far as to insist that Gerd Kùlot's review could not conceivably be about *Prince Hyppia: Math Dramatica* but had to be about *another* play, possibly a related play, perhaps an *analogous* play, perchance even a perfectly *parallel* play, peradventure a play with a similar-sounding title penned by a pair of paranoiac paradoxophobes, but in any case it was not about *his* play.

And yet, protest though he might, there was nothing at all that Y. Ted Enrustle could do about how audiences were now interpreting his beloved play's lines, because the two notions — the sauciness of certain integers and the unpennability of certain lines of theatrical dialogue — were now seen by enlightened playgoers as precisely isomorphic phenomena (every bit as isomorphic as the parallel escapades of Aurélie and Pomponette). The subtle mapping discovered by the impish Kùlot and gleefully revealed in his review made both meanings apply equally well (at least to anyone who had read and understood the review). The height of the irony was that, in the case of a few choice arithmetical remarks such as Prince Hyppia's famous

outburst, it was *easier* and *more natural* to hear them as referring to unpennable lines in plays than to hear them as referring to non-prim numbers! But Y. Ted Enrustle, despite reading Külot's review many times, apparently never quite caught on to what it was really saying.

### Analogy, Once Again, Does its Cagey Thing

Okay, okay, enough's enough. The jig's up! Let me confess. For the last several pages, I've been playing a game, talking about strangely named plays by strangely named playwrights as well as a strangely titled review by a strangely named reviewer, but the truth is (and you knew it all along, dear reader), I've *really* been talking about something totally different — to wit, the strange loop that Austrian logician Kurt Gödel (Gerd Külot) discovered and revealed inside Russell and Whitehead's *Principia Mathematica*.

"Now, now," I hear some voice protesting (but of course it's not *your* voice), "how on earth could you have *really* been talking about Whitehead and Russell and *Principia Mathematica* if the lines you wrote were not about them but about Y. Ted Enrustle and *Prince Hyppia: Math Dramatica* and such things?" Well, once again, it's all thanks to the power of analogy; it's the same game as in a *roman à clef*, where a novelist speaks, not so secretly, about people in real life by ostensibly speaking solely about fictional characters, but where savvy readers know precisely who stands for whom, thanks to analogies so compelling and so glaring that, taken in their cultural context, they cannot be missed by anyone sufficiently sophisticated.

And so we have worked our way up my ladder of examples of doubly-hearable remarks, all the way from the throwaway café blurt "This tastes awful" to the supersophisticated dramatic line "The number  $g$  is not prim". We have repeatedly seen how analogies and mappings give rise to secondary meanings that ride on the backs of primary meanings. We have seen that even primary meanings depend on unspoken mappings, and so in the end, we have seen that all meaning is mapping-mediated, which is to say, all meaning comes from analogies. This is Gödel's profound insight, exploited to the hilt in his 1931 paper, bringing the aspirations embodied in *Principia Mathematica* tumbling to the ground. I hope that for all my readers, understanding Gödel's keen insight into meaning is now a piece of cake.

### How Can an "Unpennable" Line be Penned?

Something may have troubled you when you learned that Prince Hyppia's famous line about the number  $g$  proclaims (via analogy) its own unpennability. Isn't this self-contradictory? If some line in some play is

truly unpennable, then how could the playwright have ever penned it? Or, turning this question around, how could Prince Hyppia's classic line be found in Y. Ted Enrustle's play if it never was penned at all?

A very good question indeed. But now, please recall that I defined a "pennable line" as a line that could be written by a playwright who was tacitly adhering to a set of well-established dramaturgical conventions. The concept of "pennability", in other words, implicitly referred to some particular *system of rules*. This means that an "unpennable" line, rather than being a line that could never, ever be written by anyone, would merely be a line that violated one or more of the dramaturgical conventions that most playwrights took for granted. Therefore, an unpennable line could indeed be penned — just not by someone who rigorously respected those rules.

For a strictly rule-bound playwright to pen such a line would be seen as extremely inconsistent; a churlish drama critic, ever reaching for cute new ways to snipe, might even write, "X's play is so mega-inconsistent!" And thus, perhaps it was the recognition of Y. Ted Enrustle's unexpected and bizarre "mega-inconsistency" that invariably caused audiences to gasp at Prince Hyppia's math-dramatic outburst. No wonder Gerd Külot received kudos for pointing out that a *formerly* unpennable line had been penned!

### "Not" is Not the Source of Strangeness

A reader might conclude that a strange loop necessarily involves a self-undermining or self-negating quality ("This formula is *not* provable"; "This line is *not* pennable"; "You should *not* be attending this play"). However, negation plays no essential role in strange loopiness. It's just that the strangeness becomes more pungent or humorous if the loop enjoys a self-undermining quality. Recall Escher's *Drawing Hands*. There is no negation in it — both hands are drawing. Imagine if one were erasing the other!

In this book, a loop's strangeness comes purely from the way in which a system can seem to "engulf itself" through an unexpected twisting-around, rudely violating what we had taken to be an inviolable hierarchical order. In the cases of both *Prince Hyppia: Math Dramatica* and *Principia Mathematica*, we saw that a system carefully designed to talk only about numbers and *not* to talk about itself nonetheless ineluctably winds up talking about itself in a "cagey" fashion — and it does so precisely because of the chameleonic nature of numbers, which are so rich and complex that numerical patterns have the flexibility to mirror any other kind of pattern.

Every bit as strange a loop, although perhaps a little less dramatic, would have been created if Gödel had concocted a self-affirming formula that cockily asserted of itself, "This formula is provable via the rules of

*PM*”, which to me is reminiscent of the brashness of Muhammad (“I’m the greatest”) Ali as well as of Salvador (“The great”) Dalí. Indeed, some years after Gödel, such self-affirming formulas were concocted and studied by logicians such as Martin Hugo Löb and Leon Henkin. These formulas, too, had amazing and deep properties. I therefore repeat that the *strange* loopiness resides not in the flip due to the word “not”, but in the unexpected, hierarchy-violating twisting-back involving the word “this”.

I should, however, immediately point out that a phrase such as “this formula” is nowhere to be found inside Gödel’s caggy formula — no more than the phrase “this audience” is contained in Cagey’s line “Anyone who crosses the picket line to go into Alf and Bertie’s Posh Shop is scum.” The unanticipated meaning “People in *this audience* are scum” is, rather, the inevitable outcome of a blatantly obvious analogy (or mapping) between two entirely different picket lines (one outside the theater, one on stage), and thus, by extension, between the picket-crossing members of the audience and the picket-line crossers in the play they are watching.

The preconception that an obviously suspicion-arousing word such as “this” (or “I” or “here” or “now” — “indexicals”, as they are called by philosophers — words that refer explicitly to the speaker or to something closely connected with the speaker or the message itself) is an indispensable ingredient for self-reference to arise in a system is shown by Gödel’s discovery to be a naïve illusion; instead, the strange twisting-back is a simple, natural consequence of an unexpected isomorphism between two different situations (that which is being talked about, on the one hand, and that which is doing the talking, on the other). Bertrand Russell, having made sure that all indexical notions such as “this” were absolutely excluded from his formal system, believed his handiwork to be forever immunized against the scourge of wrapping-around — but Kurt Gödel, with his fateful isomorphism, showed that such a belief was an unjustified article of faith.

### Numbers as a Representational Medium

Why did this kind of isomorphism first crop up when somebody was carefully scrutinizing *Principia Mathematica*? Why hadn’t anybody thought of such a thing before Gödel came along? It cropped up because *Principia Mathematica* is in essence about the natural numbers, and what Gödel saw was that the world of natural numbers is so rich that, given *any* pattern involving objects of any type, a set of numbers can be found that will be isomorphic to it — in other words, there are numbers that will perfectly mirror the objects and their pattern, numbers that will dance in just the way the objects in the pattern dance. Dancing the same dance is the key.

Kurt Gödel was the first person to realize and exploit the fact that the positive integers, though they might superficially seem to be very austere and isolated, in fact constitute a profoundly rich representational medium. They can mimic or mirror any kind of pattern. Like any human language, where nouns and verbs (etc.) can engage in unlimitedly complex dancing, the natural numbers too, can engage in unlimitedly complex additive and multiplicative (etc.) dancing, and can thereby “talk”, via code or analogy, about events of any sort, numerical or non-numerical. This is what I meant when I wrote, in Chapter 9, that the seeds of *PM*’s destruction were already hinted at by the seemingly innocent fact that *PM* had enough power to talk about *arbitrarily subtle properties* of whole numbers.

People of earlier eras had intuited much of this richness when they had tried to embed the nature of many diverse aspects of the world around us — stars, planets, atoms, molecules, colors, curves, notes, harmonies, melodies, and so forth — in numerical equations or other types of numerical patterns. Four centuries ago, launching this whole tendency, Galileo Galilei had famously declared, “The book of Nature is written in the language of mathematics” (a thought that must seem shocking to people who love nature but hate mathematics). And yet, despite all these centuries of highly successful mathematizations of various aspects of the world, no one before Gödel had realized that one of the domains that mathematics can model is *the doing of mathematics itself*.

The bottom line, then, is that the unanticipated self-referential twist that Gödel found lurking inside *Principia Mathematica* was a natural and inevitable outcome of the deep representational power of whole numbers. Just as it is no miracle that a video system can create a self-referential loop, but rather a kind of obvious triviality due to the power of TV cameras (or, to put it more precisely, the immensely rich representational power of very large arrays of pixels), so too it is no miracle that *Principia Mathematica* (or any other comparable system) contains self-focused sentences like Gödel’s formula, for the system of integers, exactly like a TV camera (only more so!), can “point” at any system whatsoever and can reproduce that system’s patterns perfectly on the metaphorical “screen” constituted by its set of theorems. And just as in video feedback, the swirls that result from *PM* pointing at itself have all sorts of unexpected, emergent properties that require a brand-new vocabulary to describe them.

