make sense of the neuroimaging studies, which blurred together different blips of brain activity that we can now see are nicely separated in time.

Whatever the outcome of the past-tense treasure hunt, I hope it will be emblematic of a trend in intellectual life in the coming millennium that the biologist E. O. Wilson has called consilience: the unification of the arts and sciences by an understanding of mind, brain, and human nature.[53] Regular and irregular inflection has long been mulled over by novelists and poets, dictionary writers and editors, philologists and linguists. Now this topic straight out of the humanities is being probed with the cutting-edge tools of molecular genetics and imaging of the brain. Some people fear this kind of development as a crass "reductionism" that will marginalize the humanities and plough under the richness of their subject matter, but it is far from that. Without an understanding of the contents of the mind from psychology, linguistics, and all the other disciplines they touch, neuroscientists would not know where to begin in studying the human brain, and their technologies would be expensive toys. Ultimately all knowledge is connected, and insight into a phenomenon can come from any direction, from the outcome of the Battle of Hastings to the sequence of a kinase gene.

10

A DIGITAL MIND IN AN ANALOG WORLD

The ingredients of language are words and rules. Words in the sense of memorized links between sound and meaning; rules in the sense of operations that assemble the words into combinations whose meaning can be computed from the meanings of the words and the way they are arranged. I have tried to convince you of this simple idea, and to illuminate some of the wonders of language, by exploring the ins and outs of a single curiosity in which the two ingredients may be contrasted.

Regular and irregular forms in English are the same size—one word long—and express the same ideas—past tense or plural. Yet the human mind treats them differently. Irregular forms fail to show up, and the regular pattern makes itself available, in a variety of cases that have nothing in common but a failure of access to information in memory. We have seen the regular form surface when a word is new, rare, unusual, without a standard root, or without a way for information in the root to apply to the whole word. We have seen it surface when the memories of words are freshly formed in children and when they have decayed from disease in adults.

This is an odd assortment of circumstances, some rather exotic. Surely the mind is not equipped with features designed to give rise to each of them. The simplest explanation is that regular inflection is computed by a mental operation that does not *need* access to the contents of memory: a symbol-processing rule, which attaches a suffix to any word that bears the mental symbol "verb" or "noun."

We also have seen that the power of a rule to serve as a default, stepping in when memory and analogy fail, can be observed in languages from all over the world. It does not depend on words' being frequent in the language, nor on their having distinctive sounds. Children have a sense of the kinds of words to which a rule may apply and the kinds of words to which a rule may not apply, even when they have never encountered those kinds of words before. All this suggests that a rule does not gain its power from having been pounded into the child's mind. Instead it may gain its power from the very nature of the child's mind.

I believe that regular and irregular forms show us the mental mechanisms that lie behind the two principles of language. A memory system stores and retrieves words, implementing Ferdinand de Saussure's principle of the arbitrary sign. A system of symbolic computation generates grammatical combinations of words, implementing Wilhelm von Humboldt's principle of the infinite use of finite media. Together they explain the vast expressive power of language, the ability to convey an unlimited number of new ideas.

I want to leave you with a remarkable parallel between regular and irregular inflection and something completely different. The parallel cannot be a coincidence, and it hints that the distinction between regular and irregular forms may expose even deeper principles about the nature of the mind and how it reflects the world.[1]

~

People think in categories, like "furniture," "vegetable," "grandmother," and "turtle." The categories underlie much of our vocabulary—such as the words *turtle* and *furniture*—and they underlie much of our reasoning. We are not dumbfounded by every new turtle we see; we categorize it as a "turtle" and expect it to have certain traits, like being slower than a hare and withdrawing into its shell when frightened. This means that beforehand we did not mindlessly record every turtle we had seen, like a video camera; we must have abstracted what turtles have in common. To understand mental categories is to understand much of human reasoning.

Concepts in the mind pick out categories in the world, and the simplest explanation of concepts is that they are conditions for membership in a category, a bit like definitions in a dictionary. An "odd number" is an integer that when divided by two leaves a remainder. A "bachelor" is an unmarried adult male. A

"grandmother" is the mother of a parent. A "turtle" is a reptile with a broad flattened body enclosed in a shell formed of a dorsal carapace and a ventral plastron, united at the sides.

The power of a definition is that it transcends the particulars of experience. People can recognize a new turtle when they see one, as long as it conforms to the definition. Psychologists call these categories "classical" or "Aristotelian" categories, after the Greek philosopher who emphasized logic and definitions as the basis of knowledge. For decades psychologists studied concept learning in humans and animals by presenting them with drawings of colored shapes, indicating which ones belonged to a category such as "large red square," and measuring how long it took the subjects to infer the category.[2]

All this was challenged by the Austrian philosopher Ludwig Wittgenstein in a famous passage from his *Philosophical Investigations*, a collection of ruminations published after his death in 1951:[3]

> 66. Consider for examples the proceedings we call "games." I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all?—Don't say: "There *must* be something common, or they would not be called "games"'—but *look and see* whether there is anything common to all.—For if you look at them you will not see something that is common to *all*, but similarities, relationships, and a whole series of them at that. To repeat: don't think, but look!—Look for example at board-games, with their multifarious relationships. Now pass to card-games; here you will find many correspondences with the first group, but many common features drop out, and others appear. When we pass next to ball-games, much that is common is retained, but much is lost.—Are they all "amusing"? Compare chess with noughts and crosses [tic-tac-toe]. Or is there always winning and losing, or competition between players? Think of patience [solitaire]. In ball games there is winning and losing; but when a child throws his ball at the wall and catches it again, this feature has disappeared. Look at the parts played by skill and luck; and at the difference between skill in chess and skill in tennis. Think now of games like ring-a-ring-a-roses; here is the element of amusement, but how many other characteristic features have disappeared! And we can go through the many, many other groups of games in the same way; can see how similarities crop up and disappear.
>
> And the result of this examination is: we see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail.

67. I can think of no better expression to characterize these similarities than "family resemblances"; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way.—And I shall say: "games" form a family.

And Wittgenstein did not live to see Doom, professional wrestling, or Six Degrees of Kevin Bacon.[4] As he noted, a category can be extended to embrace new cases "as in spinning a thread we twist fibre on fibre. And the strength of the thread does not reside in the fact that some one fibre runs through its whole length, but in the overlapping of many fibres."

In the 1970s the psychologist Eleanor Rosch brought Wittgenstein's ideas into psychology by showing that many human concepts picked out family resemblance categories rather than classical categories.[5]

First, with most categories it is almost impossible to find a set of membership conditions. If the definition of a "turtle" includes having a shell, what do we do with leatherbacks and other soft-bodied turtles? If a "bachelor" is an "unmarried man," does that mean the Pope is a bachelor? A "chair" needn't have legs or a seat or a back; think of that staple of the 1970s bachelor pad, the beanbag chair. Nor must it be capable of supporting a seated human—think of the Hollywood prop that disintegrates into smithereens when the bad guy smashes it over the head of the good guy. The general point is illustrated here by Opus the Penguin:

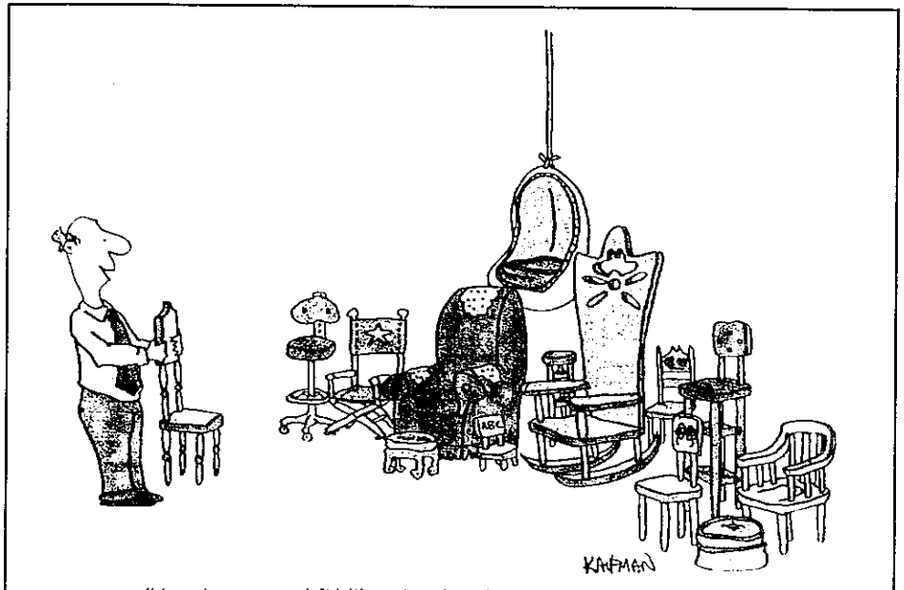
BLOOM COUNTY by Berkeley Breathed. Reprinted with the artist's permission.

Second, the members of a category are not created equal, which is what one would expect if they were admitted into the category by meeting the definition. Everyone agrees that a blue jay is somehow a better example of a bird

than a chicken or a penguin, and that an armchair is a better example of furniture than a grandfather clock. The best member of all is called the prototype, such as the sparrow for "bird" and a wrench for "tool," and it sums up the category in people's minds. Dictionaries often show a prototype next to the definition of a category. Next to the entry for *bird* you are likely to see a picture of a sparrow or a robin, not a picture of a turkey or a kiwi.

Third, the categories of the mind have fuzzy borders. People aren't quite sure whether garlic, parsley, seaweed, or edible flowers should count as vegetables, and the Reagan administration created a ruckus when it justified cutbacks in funding for school lunches by reclassifying ketchup as a vegetable. If a clamp is a tool, why not a ball of string? Is a scorpion a bug? Is a sport utility vehicle a car or a truck? Is synchronized swimming a sport?

Fourth, most of our everyday categories, and not just games, show Wittgenstein's family resemblance and crisscrossing features. Many vegetables are green, but carrots aren't; many are crunchy when raw, but spinach isn't. As for chairs, this cartoon from *The New Yorker* says it all:



"Attention, everyone! I'd like to introduce the newest member of our family."

The New Yorker Collection 1977 Jeff Kaufman from cartoonbank.com. All rights reserved.

Fifth, categories have stereotyped features: traits that everyone associates with the category, even if they have nothing to do with the criteria for membership. When people think of a grandmother, they think of gray hair and chicken soup, not of a node in a genealogical tree.

Many experiments have confirmed that everyday concepts act like family resemblance categories.[6] People are comfortable with the very idea that categories have better and worse members: They have no trouble rating the "goodness" of the members of a category on a scale of 1 (best) to 7 (worst). For example, they give a robin an average rating of 1.1 on the bird scale, and a chicken a rating of 3.8. Football was judged a fine example of a sport, earning a rating of 1.2; wrestling was a so-so example, eking out a 4.7. A carrot is a vegetable par excellence (1.1), but parsley is a more dubious instance (3.8). Murder is an excellent crime at 1.0, but vagrancy is not so good at 5.3. The ratings of different people agree closely.

When people are shown pictures of objects and asked to press a button if the object belongs to a named category, they press the "fruit" button more quickly to a picture of an apple than to a picture of a watermelon. That suggests that the category "fruit" is more easily evoked in people's minds by the apple. Rosch asked people to make up sentences with category words such as "bird." Typical responses were "I heard a bird twittering outside my window" and "Three birds sat on the branch of the tree." Then she replaced the word "bird" by various species: sparrow, penguin, eagle, ostrich. The absurdity of "I heard a penguin twittering outside my window" and "Three ostriches sat on the branch of the tree" shows that it must have been prototypical birds that had popped into the subjects' minds. Children have similar intuitions: When they first learn a word, they use it with prototypical members of a category: *bird* is used with sparrows, *vegetable* with carrots or celery.

It's also easy to show in the lab that people are fuzzy about borderline cases. The psychologists Michael McCloskey and Sam Glucksberg asked subjects to give true-or-false verdicts on category membership. Everyone agreed that cancer is a disease, that apples are fruit, and that flies are insects. But when it came to deciding whether stroke is a disease, a pumpkin is a fruit, or a leech is an insect, half the subjects went one way and half went the other—and when they were asked again a month later, many changed their minds.[7]

Does this mean that people's heads are stuffed with fuzz and that classical categories are fictions? Surely not. People can learn categories with clean definitions, crisp edges, and no family resemblance, such as "odd number." They can learn that a dolphin is not a fish, though it has a strong family resemblance

to the fishes, and that a seahorse is a fish, though it looks more like a little horse. They can understand that Tina Turner is a grandmother, though she lacks all the usual traits, and that my childless great-aunt Bella was not a grandmother, though she had gray hair and made a mean chicken soup. Though people refer to women in their third trimester as "very pregnant," they also understand what it means when parents say to their daughters, "You can't be just a little bit pregnant."[8]

The psychologists Sharon Armstrong and Henry and Lila Gleitman replicated Rosch's experiments using the most classical, Aristotelian categories they could find, "odd number" and "woman." The subjects rated "7" as an excellent example of an odd number, and "447" as not such a good example; they thought that a "housewife" was an excellent example of a woman, and a "policewoman" not such a great example. The same gradations emerged in their real-time mental processes: They pushed an "odd number" button more quickly when "3" flashed on the screen than when "2,643" did. Surely those students would not have made it into the prestigious University of Pennsylvania if they really thought that numbers could be more or less odd, and indeed in a questionnaire they averred that a number was either even or odd, with no in-between cases. So they must have been capable of turning their fuzziness on and off. Family resemblance categories are real, but so are classical categories; they live side by side in people's minds, as two ways of construing the world.[9]

~

What does this have to do with regular and irregular verbs? The psychologist Dan Slobin and the linguist Joan Bybee were the first to point out that classes of irregular verbs with similar past-tense forms, such as *sing–sang, ring–rang, drink–drank* and *bind–bound, find–found, grind–ground,* are just like Wittgenstein's family resemblance categories.[10] All five of their distinguishing traits can be found in the irregulars.

First, despite the contortions of centuries of language scholars, no one has been able to craft a set of rules that properly pick out the different kinds of irregular verbs. As Mark Twain said of a German grammar book, there are more exceptions to a rule than there are instances of it. For example in English the largest family of irregular verbs are no-changers like *rid–rid, cut–cut,* and *set–set.* They all end in *t* or *d,* but there is no hope of lassoing the family with a rule stipulating that verbs ending in *t* or *d* belong to it. Next to *hit–hit, slit–slit, split–split,* and *quit–quit* we find regular *flit–flitted, twit–twitted,* and *pit–*

pitted. Near *let* and *set* we find regular *fret, sweat*, and *whet*. Beside *cut* and *shut* we find *butt, jut*, and *strut*. Adjacent to *hurt* we find *blurt* and *spurt*; near *burst* we find regular *bust*. All the other classes of irregulars have rule defeaters too, as we saw in chapter 5.

Second, in every irregular family some members are more equal than others. *Hit* and *split* are full-fledged no-changers, but in the minds of Americans *spit–spit* and *forbid–forbid* are so-so. Ditto for the other families:

Good Examples	Poor Examples
bleed–bled, feed–fed	*plead–pled, speed–sped*
burn–burnt, bend–bent	*learn–learnt, lend–lent, rend–rent*
deal–dealt, feel–felt, mean–meant	*kneel–knelt, dream–dreamt*
freeze–froze, speak–spoke	*weave–wove, heave–hove*
get–got, forget–forgot	*beget–begot, tread–trod*
write–wrote, drive–drove, ride–rode	*smite–smote, strive–strove, stride–strode*

Many of the classes have a prototype or best kind of member. For the *ing–ung* family it is verbs that fit the pattern *s-consonant-consonant-ing*, such as *string*. Bybee discovered that people are most tempted to grant an irregular form to a made-up verb when the verb matches the prototype, as in *spling–splung* and *skring–skrung*.

Third, in the halo around the poor relations in an irregular family there are verbs so poor that no one knows whether they belong in the family at all:

He has *stridden* around the park three times.
They seem to have *striven* to baffle their readers.
I don't know how she *bore* the guy.
I *forwent* the pleasure of grading papers last night.
The mice *throve* in the compost.

Fourth, the members of irregular families resemble each other in crisscrossing ways, rather than by sharing any trait. Take the second-biggest family of irregular verbs, the *ring–rang–rung* family, which change *ĭ* to *ă* or *ŭ*. Most of the members end with the consonant *ng*, which is velar (pronounced at the velum or soft palate) and nasal (pronounced through the nose): *shrink, sink, stink,*

cling, fling, sling, slink, sting, string, swing, and *wring*. That screams for a rule that states, "Change *ing* to *ung*." But the rule runs afoul of the crisscrossing resemblances, as we saw in chapter 4. Some family members end in a consonant that is velar but not nasal: *stick, dig, sneak*, and *strike*. Others end in a vowel that is nasal but not velar: *win, spin, swim*, and *begin*. The rule would miss them all.

The other families of irregulars criss and cross as well. *Blow–blew, grow–grew*, and *throw–threw* begin with a cluster of consonants and end in *ō*. *Draw, fly*, and *slay* have the consonants but not the vowel, and *know* has the vowel but not the consonants. Incidentally, the spelling of *know*, which once reflected its pronunciation, shows that the word *used* to have a consonant cluster, like its relatives, before English speakers stopped pronouncing the *k* in words like *knee, knife, knob*, and *knuckle*. The *ow–ew* class started out neat and became ragged, a fact to which we will return.

Fifth, irregular families have stereotyped features that run in the family but play no role in defining the past-tense form. Take the verbs that change *d* to *t*, such as *bend–bent*. In principle any verb ending in *d* could see it replaced by *t*: Our language could have given us *sled–slet, fold–folt*, and so on. In reality almost all of these verbs end in *-end*: *lend–lent, send–sent, spend–spent, bend–bent*. Similarly, one can imagine a language in which any *ā* could become *o͝o* (as in *foot*), but in English the *ā*s that do give way to *o͝o* are preceded by a tongue-tip consonant and followed by *k*: *take–took, shake–shook, forsake–forsook*.

Clusters of irregular verbs pass all five tests of Wittgenstein's family resemblance categories. In his book *Women, Fire, and Dangerous Things* (a family resemblance category in an Australian aboriginal language), the linguist George Lakoff called attention to the fuzziness that lies at the heart of that traditional bastion of rules, grammar. He cited irregular verbs as the ultimate proof of the bankruptcy of the two-thousand-year-old Aristotelian tradition in Western thought that seeks precise definitions for everything in sight.[11]

But Lakoff did not notice that right next door to the irregulars are the *regular* verbs, and they pass all the tests of *classical* categories. Other than verbs with an irregular form in memory, all verbs are members of the regular family in equal standing, simply by meeting the criterion "is a verb." As we have seen, regular verbs can have any sound: sounds that are strange in English, as in *ploamphed, oinked*, and *out-Gorbachev'd*; sounds that are already associated with irregular verbs, as in *high-sticked* and *flied out*; and sounds that have rarely or never been heard before, as in *Borked* and *anastomosed*. People find

ploamphed to be as good a past tense of *ploamph* as *plipped* is of *plip*, and they produce and approve the past-tense forms of rare verbs like *balk* as readily as they do with common verbs like *walk*. The regular verbs do not fall into clusters, have no stereotypes, no family resemblance, and aside from occasional interference from irregular verbs, no fuzzy examples.

Why on earth should irregular verbs act like games and furniture and vegetables, and regular verbs act like grandmothers and odd numbers? Are we seeing the outward signs of some deep common cause, or is it all a coincidence, worthy of attention only from conspiracy buffs? I believe there is something beneath the similarities, and that the facts of regularity and irregularity offer glimmers of insight into the nature of our conceptual categories.[12] These facts shed light on the mental machinery that computes our conceptual categories and on the things in the world that our conceptual categories are good at picking out.

~

Regular and irregular forms coexist but require different computational mechanisms: symbol combination for regular forms, associative memory for irregular forms. The same may be true for classical and family resemblance categories.

Before Rumelhart and McClelland built their pattern associator for the past tense, they built one for conceptual categories. It learned concepts like "dog," "cat," and "bagel" by picking up associations among the perceptual features (furry, four-legged, and so on) that tend to co-occur in them.[13] For example, the concept "cat" was implicit in a pattern of strong connections among units that stand for the typical traits of cats, such as whiskers, meowing, and pointy ears. The pattern associator reproduced most of the signatures of family resemblance categories that Rosch had demonstrated in human beings, such as responding to prototypical cats more strongly than to atypical cats. Many subsequent models have had similar success.[14] That is because a family resemblance category is held together by crisscrossing traits, and a pattern associator is a gadget for learning how traits crisscross.

But just as pattern associators for the past tense are good at some things and not so good at others, so too are pattern associators for concepts. A model that is good at picking up stereotypes is apt to project the stereotype onto atypical objects. One model, for example, when taught that a plate had broken, ignored the teacher and concluded that the object was either a window or a vase, because all the broken objects in its training set were windows or vases.[15] Another, when told that an office had drapes, concluded that it wasn't an office, because all the offices in its training lacked drapes.[16] Gary Marcus has shown that standard pattern associators cannot generalize from "a skunk has skunk babies," "a cat has cat babies," and "a bear has bear babies" to "a greeble has greeble babies" (where a greeble is a newly encountered animal), because they lack a variable, "X," that would allow them to learn that "an X has X babies."[17] These failures are reminiscent of the past tense models' habit of turning out strange blends, or nothing at all, when fed rare or unusual words.

The facts about verbs and the facts about concepts converge to suggest that the human mind is a hybrid system, learning fuzzy associations and crisp rules in different subsystems. Most of the recent models of human categorization in cognitive psychology (which are designed to capture people's speed and accuracy when learning artificial categories in the lab) are built out of two parts: a pattern associator for categories based on families of similar exemplars, and a rule selector for categories based on rules. The psychologists were forced to these hybrid models because with some categories subjects quickly figure out a rule (such as "rectangles that are taller than they are wide"), whereas with other categories subjects go by their gut feelings, memorizing some of the examples and classifying the new ones according to how similar they are to the memorized ones. No model that uses a single mechanism to capture people's behavior with every kind of category does as well as the hybrid models.[18] Some modelers even link the rule system to the frontal cortex and the exemplar-based system to the temporal and posterior cortex, much as we did for rules and words in the preceding chapter.

~

Why do we have these two ways of knowing? It is unlikely that natural selection equipped us with mental machinery that is completely out of synch with the world in which we live. Might the difference between classical and family resemblance categories reflect a difference between two kinds of things in the world, or at least two ways of reasoning about things in the world? In the case of the past-tense system, we know the ancestry and logic of the verbs in considerable detail. Perhaps they have something to teach us about the different kinds of conceptual categories.

Irregular forms are relics of history. They fall into families because originally they were generated in matched sets by rules, but the rules died long ago and

the families have been disintegrating ever since. Vowels drift, consonants get swallowed, words lose their popularity, dialects break apart or coalesce. After centuries or millennia irregular forms are no longer the orderly outputs of a rule, nor are they a list of unrelated sounds; they are a family resemblance category. A clear example is the verb *know–knew*, which used to have a consonant cluster like its siblings *grow–grew, blow–blew*, and *throw–threw*, but then lost its first consonant, messing up the class.

Children are born into a linguistic world that throws the members of a family resemblance category at them, and they cope quite well. With their pattern-loving memories they reproduce most of their parents' irregular verbs. Occasionally they lose an irregular like *chide–chid* or *seem–sempt*, but occasionally they add one like *kneel–knelt* or *sneak–snuck*. The adopted word shares some traits with its new family, because it was that similarity that attracted it in the first place. But each adoptee brings some unique traits in with it, so the class remains ragged. The next generation also finds itself with a family resemblance category to commit to memory.

Regular past-tense forms, in contrast, have no history. In fact they barely have an existence. Only the past-tense *rule* exists. Children don't have to cope with learning the quirks of regular forms because they don't have to learn regular forms at all. The rule creates them when they are needed, and then they can be thrown away, because the rule is always around to create them again the next time. Now, that is an exaggeration—children have to remember a few regular forms to learn the rule to begin with, and adults certainly do remember many regular forms alongside the rule. Yet once the rule is acquired, the forms don't *need* to exist for speakers to use and understand them. The category of regular forms is not a real category but a virtual category: the list of forms that *would* be created if the rule were allowed to work its way through all the verbs in a person's vocabulary. Children never see the category; what they learn is not a class of regular forms left behind by previous speakers but a rule that matches the rule in other people's heads.

The past-tense rule itself would hardly be worth the trouble were it not part of the magnificent *system* of rules we call grammar. Children are wired to learn that system, which allows us all to convey an infinite number of brand-new thoughts. The category of regular forms is a by-product of the rule system.

The two kinds of conceptual categories, I think, harmonize with two kinds of things in the world in the same way that regular and irregular verbs harmonize with two kinds of things in the minds of other speakers.

To see this we have to begin at the beginning. Why does the mind even *have* categories like "birds" and "games"? No two inhabitants of the world are identical, and one can imagine a mind that treated every object as a unique individual, just as we treat our friends as unique individuals. In fact we don't have to imagine such a mind; Jorge Luis Borges has imagined him for us, in his story "Funes the Memorious":

> We, at one glance, can perceive three glasses on a table; Funes, all the leaves and tendrils and fruit that make up a grape vine. He knew by heart the forms of the southern clouds at dawn on the 30th of April, 1882, and could compare them in his memory with the mottled streaks on a book in Spanish binding he had only seen once and with the outlines of the foam raised by an oar in the Río Negro the night before the Quebracho uprising. . . . A circle drawn on a blackboard, a right triangle, a lozenge—all these are forms we can fully and intuitively grasp; Ireneo could do the same with the stormy mane of a pony, with a herd of cattle on a hill, with the changing fire and its innumerable ashes, with the many faces of a dead man throughout a long wake. . . .
>
> Not only was it difficult for him to comprehend that the generic symbol *dog* embraces so many unlike individuals of diverse size and form; it bothered him that the dog at three fourteen (seen from the side) should have the same name as the dog at three fifteen (seen from the front). His own face in the mirror, his own hands, surprised him every time he saw them. Swift relates that the emperor of Lilliput could discern the movement of the minute hand; Funes could continuously discern the tranquil advances of corruption, of decay, of fatigue. He could note the progress of death, of dampness.[19]

Why are we not like Funes the Memorious? Are we just an anal retentive species that likes to put things into pigeonholes for the sheer orderliness of it all? And if we are, how do we decide on the pigeonholes? There are a frightful number of ways to sort objects into categories—alphabetically, in pairs, according to height, and so on. Why "birds"?

The answer is that people form categories that give them an advantage in reasoning about the world by allowing them to make good predictions about aspects of an object they have not directly seen. We cannot bring every object home and put it under a microscope or send tissue samples out for lab testing. We have to observe a few traits that the object wears on its sleeve and infer the

traits that we cannot see directly. Good categories let us do that. If Tweety has feathers and a beak, Tweety is a bird; if Tweety is a bird, Tweety is warm-blooded, can fly, and has hollow bones. Bad categories do not: If we knew only that Tweety's name begins with a "T," nothing of interest would follow.[20]

These inferences work only if the world is properly structured. If a capricious god had assembled every object with a unique, random combination of traits, like the numbers on a lottery ticket, inference would be impossible. The blood of a feathered friend would be cold as often as hot, its bones solid as often as hollow. Luckily for us, we don't live in that world. We live in a lawful world in which traits tend to hang together in the same way in many objects.

Our mental categories are useful because they reflect the lawfulness of the world. In theory, laws could be apprehended in different ways. At one extreme one could extract the underlying laws directly and use them in chains of deduction. An example is using the laws of solid geometry and thermodynamics to predict that small animals lose heat faster than large ones, because heat is lost at surfaces and small things have a greater ratio of surface area to volume than large things. At the other extreme we can assemble an enormous database by measuring every trait of every object we can find, and when faced with a new object, find the closest old object and predict it is similar. If we learn that sparrows lose heat quickly, we guess that starlings do too.

The first method seems powerful and insightful, the second one mindless and drudgelike. But often mortal knowers have no choice but to use the second. As the poet John Ciardi wrote,

> Who could believe an ant in theory?
> A giraffe in blueprint?
> Ten thousand doctors of what's possible
> Could reason half the jungle out of being.

Many things we find around us could not be deduced by any body of laws, because they are shaped by myriad events of history no longer visible to us.

Take birds. In the course of evolution a species begins in a population of interbreeding organisms adapting to an ecological niche. Natural selection "engineers" the organisms to compete well in that environment, and sexual reproduction homogenizes them. If we could go back and look at the last common ancestors of the birds, they would be as similar as a single species is today. They would be genetically similar because they descended from a common set

of ancestors, bred with one another, and underwent natural selection for traits such as wings and a streamlined shape. But that uniformity did not last for long. Their descendants went off in different directions, some becoming nocturnal like the kiwi, some taking to the sea like penguins, some growing large like ostriches. The aftermath of this radiation is a family resemblance category. All birds have common traits such as beaks and wings and feathers because they inherited them from their common ancestor. But other traits crisscross, rather than running throughout the class, because each species has a unique history in which some traits were lost and others acquired.

And this brings us to the parallel between irregular verbs and family resemblance categories. Irregular families were once generated by rules but then accumulated idiosyncrasies, and now they must be memorized individually. Classes of animal species were once adapted to a single niche but then dispersed and accumulated idiosyncrasies, and now each species must be learned about through observation. In each case the surviving similarities in the family members are too useful to ignore, and the memory system extracts the patterns rather than filing each item in a separate slot in memory. The patterns determine the better and worse members, and they allow a knower to guess that newly encountered similar items belong to the family.

A comparison of the history of words and the history of species may strike you as far-fetched, but it has a distinguished background. Darwin himself illustrated his key idea—that the similarities and differences among organisms could be explained by their family history—by analogy to how words change in languages:

> The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously parallel. . . . We find in distinct languages striking homologies due to community of descent, and analogies due to a similar process of formation. . . . The frequent presence of rudiments, both in languages and in species, is still more remarkable. The letter *m* in the word *am*, means I; so that in the expression *I am*, a superfluous and useless rudiment has been retained. In the spelling also of words, letters often remain as the rudiments of ancient forms of pronunciation. . . . We see variability in every tongue, and new words are continually cropping up; but as there is a limit to the powers of the memory, single words, like whole languages, gradually become extinct. As Max Muller has well remarked: "A struggle for life is constantly going on amongst the words and grammatical forms in each language. The better, the shorter, the easier forms are constantly gaining the upper hand,

and they owe their success to their own inherent virtue." To these more important causes of the survival of certain words, mere novelty and fashion may be added; for there is in the mind of man a strong love for slight changes in all things. The survival or preservation of certain favoured words in the struggle for existence is natural selection.[21]

The analogy lives today in modern biology and linguistics. Similar statistical techniques are used to find the best groupings of organisms and to find the best groupings of languages, based on the co-occurrences of their traits. When biologists are unsure of which species to lump together in a genus or family, they sometimes take hundreds of measurements of animals' parts and feed them into an algorithm that finds the best categories in which to lump them.[22] Similarly, when linguists are unsure of which languages to lump together in a family, they sometimes feed hundreds of sets of cognate words into an algorithm that finds the best families in which to lump them. These algorithms are not literally pattern associator memories, but they rely on the same principle: Entities that share many traits probably come from the same category and should be treated alike.[23]

Not all family resemblance categories start off in lockstep and then diversify, but probably all of them are governed by hidden laws that make them similar and historical contingencies that make them different. Today's chairs did not descend from some ancestral ur-chair, so what makes them similar? It is that they must hold up a human bottom, and that forces most of them to have a stable, accessible, elevated, weight-bearing platform. At the same time chairs differ because of local variations in styles, tastes, materials, and expertise. Games are similar because they are meant to amuse, and they differ because of countless historical and local circumstances—the invention of playing cards, the invention of the computer, the availability of ice, grass, or water, the locals' taste for spectacle, violence, or brain work.

If we evolved a taste for family resemblance categories because they really do exist in the world as a product of history, why did we also evolve a taste for classical categories? I think it is because classical categories are by-products of rules in the mind that allow us to exploit laws in the world. The rules thereby allow us to deduce predictions about how things in the world work. Classical categories are not free-floating definitions, useful only for pigeonholing things. They always are part of a *system* of interlocking rules that churn out handy deductions or computations. Just as regular verbs are products of a rule system (grammar), classical categories are products of their own rule systems. Odd numbers belong to arithmetic, triangles to geometry, grandmothers to kinship,

dolphins to biological taxonomy, pregnancy to physiology, presidents to law. Each system allows a person to deduce unobserved traits from observable ones, not by remembering that they co-occurred but by cranking through a chain of implications. Using the rules of arithmetic, one can deduce that a set of forty-three objects cannot be divided into two equal parts. Using the rules of kinship, one can deduce that one's grandmother is the daughter of one's great-grandparents. Using the laws of physiology, one can deduce that a pregnant woman will become a parent, unless she has a miscarriage, abortion, or stillbirth. Using the laws of zoology, one can deduce that dolphins suckle their young and periodically surface to breathe. Using the laws of the land, one can deduce that the President of the United States was born in the United States more than thirty-five years ago.

Since these rule systems are, like grammar, combinatorial and recursive, they allow us to reason about an unlimited range of cases, often far from our experience. The laws of kinship allow us to say something about our family tree not just a hundred years ago but a hundred thousand or a million years ago. We can predict that if the United States still exists in 2804, there will be a presidential election that year.

When we use a system of rules, we have to turn off the family resemblance system, just as we seal off our memory for similar verbs when applying a rule to a verb that has to be regular. Within our systems of reasoning about kinship and law, a grandmother doesn't have to be grandmotherly nor a president presidential. It doesn't matter that a dolphin looks like a fish, or that the sides of a real-world triangle are not infinitesimally thin or perfectly straight. The human mind can think in *idealizations*, reducing an object to an austere description of the variables manipulated by the rule system, such as generation and gender in the case of kinship or the outcome of the electoral process in the case of law.

Science in particular depends on the mind's ability to think in idealizations, such as point masses, frictionless planes, perfect vacuums, and randomly interbreeding populations of organisms. The laws of science can be categorical statements uncluttered by the grubby details of the objects they refer to, and that allows them to be chained together in long inferences that lead to counterintuitive but correct conclusions—for example, that heat consists of moving molecules and that people and fish are cousins. That would never happen if the only form of human reasoning were the habit of generalizing similar traits to similar objects.

Of course not all people know formal science, but everyone knows a folk science (often blended with religion), in which the world and its parts are explained by elaborate interactions of hidden forces, traits, and essences. *Homo*

sapiens has been said to occupy the "cognitive niche" in nature:[24] We use knowledge of cause and effect to think up novel, complex sequences of behavior that defeat the defenses of plants and animals. People in all cultures, including hunter-gatherers whose lifestyle resembles that of our evolutionary ancestors, transcend their experience of concrete events, dig beneath appearances to ferret out laws, and combine these laws in their mind's eye to manipulate the world to their advantage. They assemble complicated traps, snares, and weapons. They recognize a few scratches on the ground as the tracks of an animal of a certain size, species, and condition, and predict its destination so they can ambush it. They remember a flower in the spring and return to it in the fall to dig up the underground tuber that has invisibly grown in the interim. They extract juices and powders from plants and animals and turn them into medicines and poisons.[25] None of these acts of creation would be possible if the mind simply remembered objects and expected similar ones to behave similarly. They depend on abstract, combinatorial reasoning, of the kind made possible by rules and variables.

Some rule systems help us deal with the material world, but many help us deal with one another. The problem with fuzzy boundaries is that people can claim to see the edges of the boundary in different places. A child doesn't go to bed one evening and wake up as an adult. At some point the child may deem himself mature enough to drink or drive, whereas others may not want to take the chance of letting him. Love grows and deepens with time, but on a given day one lover may see the relationship as having ripened to a lifelong exclusive commitment, while the other lover—and interested third parties—may have a different opinion. Several people may be wise and powerful enough to merit the leadership of a group, but when a decision has to be made for the whole group, only one voice can prevail. People stave off border disputes around socially touchy categories by implementing rules that artificially sharpen the borders. They make up conditions for adulthood, marriage, and rank, complete with rites of passage that make entry into the categories instantaneous.

We have seen that much of the richness of language comes from the tension between words and rules. In the same way, much of the richness of the public sphere of life comes from tensions between family resemblance categories built from experience and the classical categories defined by science, law, or custom. Family resemblance categories resonate with common sense, but leave us groping when faced with something that is neither fish nor fowl. Classical categories offer neat divisions, but are bound to seem legalistic, pedantic,

or abstruse. Is a fertilized ovum with a full complement of human DNA a person? What about a cell scraped from a cheek that has a full complement of human DNA and which can be cloned into a person? In surrogate births, who is the real mother: the woman who donated the egg or the woman who bore the child? Is the perpetrator of a crime innocent if he is freed on a technicality? Should a difficult court case be resolved by appealing to the most similar precedent or by appealing to constitutional principle? In 1999 President Bill Clinton was impeached for perjury after he denied having sex with his intern, Monica Lewinsky, despite their having engaged in fellatio. Clinton had treated "sex" as a classical category—a list of anatomical configurations stipulated by the law—and his adversaries treated it as a family resemblance category.

We have digital minds in an analog world. More accurately, a part of our minds is digital. We remember familiar entities and their graded, crisscrossing traits, but we also generate novel mental products by reckoning with rules. It is surely no coincidence that the species that invented numbers, ranks, kinship terms, life stages, legal and illegal acts, and scientific theories also invented grammatical sentences and regular past-tense forms. Words and rules give rise to the vast expressive power of language, allowing us to share the fruits of the vast creative power of thought.