

2

The Life Spans of Genres

Although debates about genre are as old as literary criticism itself, the term is still only loosely defined. Part of the reason is that genres look like different things at different points in the life of a text. Scholars of rhetoric, interested in the creation of works, tend to interpret genres as communicative actions.¹ Sociologists often emphasize communities of readers.² Literary scholars, for their part, are preoccupied with the patterning of the texts themselves. Of course, all of these aspects of genre are connected. But it doesn't follow that the histories of production, reception, and form are precisely the same. Does the persistence of science fiction fandom, for instance, imply that a single textual pattern has united works of science fiction themselves, from the pulp era to the present? Many scholars are skeptical. Some argue that "there is no such thing as science fiction" in the texts themselves: the history of a genre is only the history of a community of readers, who may use the term "science fiction" to describe radically different things in different decades.³

1. Carolyn Miller, "Genre as Social Action," *Quarterly Journal of Speech* 70 (1984): 151–67.

2. Pierre Bourdieu, *Distinction: A Social Critique of the Judgment of Taste* (Cambridge, MA: Harvard University Press, 1984); Paul DiMaggio, "Classification in Art," *American Sociological Review* 52 (1987): 440–55.

3. Mark Bould and Sherryl Vint, "There is No Such Thing as Science Fiction," in *Reading Science Fiction*, ed. James Gunn, Marleen Barr, and Matthew Candelaria (New York: Palgrave, 2009), 43.

Science fiction is not the only genre raising problems of this kind. Some scholars trace contemporary detective fiction, for instance, to a durable model established by Edgar Allan Poe in the 1840s. But Franco Moretti has suggested that genres rarely survive longer than a single generation. In that case, the superficial continuity of the term "detective fiction" might cover a series of different forms (the Holmesian "case," the closed circle of country-house suspects, the crime thriller), linked by inheritance more than direct similarity. Maurizio Ascari pushes this theory of flux even farther, arguing that "all genres change ceaselessly," and crime fiction no longer has much to do with Poe at all.⁴

These questions about the life spans of genres remain perplexing because they are deeply perspectival. Our disagreement is not fundamentally about texts or about the history of reception but about the possibility of comparing different historical vantage points. Everyone admits, for instance, that scientific romances by Mary Shelley and Jules Verne differ substantially from the twentieth-century phenomenon that Hugo Gernsback named "science fiction." But everyone also admits that contemporary readers of science fiction often perceive nineteenth-century romances as examples of the genre they enjoy. The controversial question is whether these perceived continuities are simply retrospective projections. It remains a difficult question because history only moves in one direction. We have no way to avoid thinking of science fiction when we read Jules Verne, so it is hard to know whether to trust our perception of continuity. Readers in 1865, conversely, had no opportunity to think of *science fiction*, since the phrase didn't yet exist. We might like to travel back to 1865, offer readers our concept of science fiction, and ask how well it describes Verne's position in their own world. But of course, if we had a time machine, we wouldn't need to define science fiction at all.

4. Franco Moretti, *Graphs, Maps, Trees: Abstract Models for Literary History* (New York: Verso, 2005), 18–22; Maurizio Ascari, "The Dangers of Distant Reading: Reassessing Moretti's Approach to Literary Genres," *Genre* 47, no. 1 (2014): 15.

In this chapter, predictive models will become not quite a time machine but something almost as useful: a memory-wiping flashbulb that allows us to strategically erase our knowledge of the future or past as needed. The computer knows nothing about literary history: it models only the evidence we give it. This useful blindness will allow us to provisionally bracket twentieth-century science fiction and to model Verne purely by contrasting him to his nineteenth-century contemporaries. Then we can compare those models to models of the twentieth-century genre and see how closely their predictions align. In the pages that follow, I will call this method “perspectival modeling.” This chapter will often literally be modeling the perspectives of different groups of readers, but the phrase could also apply more broadly to any strategy that gets leverage on cultural history by comparing models trained on different subsets of the evidence. The end of this chapter will assess the pace of change, for instance, by comparing models trained on different periods.⁵

The strength of this method comes from something that might appear to be its weakness. A supervised model learns only from the examples we give it. This means that its predictions have a purely relative significance. The model cannot measure any universal dimension of language; it just indicates whether a given text more closely resembles the examples in group A or group B. At first, the relative character of this evidence may appear to undermine the whole point of quantification. We often imagine numbers as an objective yardstick—something like sentence length or linguistic entropy that can become a fixed reference point for the flux of history. A supervised model, by contrast, provides only a circular sort of evidence: it describes one set of texts through their relation to another.

But then, circularity has long been recognized as the inherent condition of historical interpretation.⁶ Events and artifacts ac-

5. The germ of insight for this method came from Michael Witmore, “The Time Problem: Rigid Classifiers, Classifier Postmarks,” *Wine Dark Sea*, April 16, 2012, <http://winedarksea.org/?p=1507>.

6. Wilhelm Dilthey, *The Formation of the Historical World in the Human Sciences* (1910; Princeton, NJ: Princeton University Press, 2002), 280–81.

quire meaning only in relation to other parts of history: they are always interpreted in relation to their own historical context, as well as other contexts implicitly brought to bear by an interpreter. So a method that provides relative and contextual description may be exactly what we need to orient ourselves in this space. When researchers try to step outside history to find a fixed metric like sentence length, they often end up creating a measurement whose significance is hard to define. Human significance is created only by human history. For that reason, perspectival modeling will become this book’s central source of leverage on the past—both in this chapter’s arguments about genre and in subsequent arguments about gender and literary prestige.

This mode of analysis is admittedly more complex than simply measuring linguistic features of the text. But it becomes complex in order to address a problem fundamental to the humanities. Statistical models can embrace the contextual and perspectival character of historical interpretation, while giving scholars more control over the contexts and perspectives they provisionally adopt. This will never allow us to entirely escape our own historical situation. Although the models in this chapter rely on books labeled by many different hands across hundreds of years, they are still produced in the twenty-first century, with twenty-first-century methods. (As I have said, a model is not literally a time machine.) But perspectival modeling does at least allow us to represent the perspectives of other eras, in a form solid enough to allow one perspective to be compared rigorously to another.

The models of genre created in this chapter will rely on textual evidence. But that doesn’t force us to assume that genres are fundamentally textual patterns. As I began by noting, genres wear several different faces: they are practices of literary production, horizons of readerly expectation, and textual templates, all at once. If I had to choose one of those three faces as primary, I would lean toward emphasizing readers’ expectations. So this chapter will initially define a “genre” as a group of books recognized by some specific, historically situated group of readers—whether they are nineteenth-century reviewers arguing about sensation novels or twentieth-century librarians shelving books

under “Detective Fiction.” Textual similarity will enter this chapter’s argument only because it turns out that texts can in fact be used to predict readers’ responses. A model trained on a sample of texts that reviewers labeled “detective fiction” can identify other books that the same reviewers labeled the same way.⁷ In other words, text will function here as a medium, registering definitions of genre that are in themselves fundamentally social. One advantage of the textual medium is that it allows us to compare definitions formed in different historical contexts and expressed through groups of illustrative examples that may not overlap.

What Story Should We Expect to Find?

Since literary scholars adhere to radically different theories of genre, we might expect several different outcomes for this inquiry. If Mark Bould and Sherryl Vint are right that “there is no such *thing*” as genre in the text itself—only a contested set of expectations defined and redefined by marketing—then we might expect textual models of genre simply to fail. If Franco Moretti is right that genre is a generational phenomenon, we might expect textual models to be stable only for twenty or thirty years. Science fiction, in particular, seems likely to be mutable, since submarines and ray guns are constantly replaced by new wonders of tomorrow. It is not self-evident that the textual patterns in science fiction would persist for more than a few decades.

On the other hand, we might just as reasonably expect to find a different pattern. With genre-specific pulp magazines (like *Amazing Stories*) and imprints (like Harlequin), twentieth-century genres achieved an institutional solidity that had rarely

7. This chapter’s work on genre was inspired by several trailblazing experiments. Brett Kessler, Geoffrey Nunberg, and Hinrich Schütze, “Automatic Detection of Text Genre,” *Proceedings of the Eighth Conference of the European Chapter of the Association for Computational Linguistics* (1997): 32–38; Sarah Allison, Ryan Heuser, Matthew Jockers, Franco Moretti, and Michael Witmore, “Quantitative Formalism: An Experiment,” Stanford Literary Lab Pamphlet 1, January 15, 2011, <https://litlab.stanford.edu/LiteraryLabPamphlet1.pdf>; Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History* (Urbana: University of Illinois Press, 2013), 67–92.

been seen before. Histories of genre often emphasize the gradual consolidation of genre conventions in the first half of the twentieth century. Scholars of the detective story point to Ronald Knox’s so-called Decalogue (1929) of rules for writers as a crystallizing moment.⁸ For other genres the process of consolidation is thought to have taken even longer. Gary K. Wolfe suggests that “the science fiction novel persistently failed to cohere as a genre” until Pocket Books gave it institutional form in the 1940s.⁹ If this account of literary history is correct, we might expect to find not a succession of generational phases but a steady hardening of boundaries, producing genres that are much more clearly distinct by the middle of the twentieth century than at its outset. Frankly, this is what I expected to find when I began the project.

To investigate these questions, I have gathered lists of titles assigned to a genre in eighteen different sites of reception.¹⁰ Some of these lists reflect recent scholarly opinion; some were shaped by nineteenth-century reviewers; others reflect the practices of many different library catalogers over a long period of time. Although each list defines its object slightly differently, they can be loosely arranged around three master categories whose coherence I propose to test: detective fiction (or “mystery” or “crime” or “Newgate” fiction), science fiction (also defined in a variety of ways), and the Gothic. (It is debatable whether the Gothic, writ large, is a genre at all—but that’s what makes it an interesting case.) I also collected texts corresponding to these titles, relying on the Chicago Text Lab and HathiTrust Digital Library as sources. By comparing groups of texts associated with different communities of reception and segments of the timeline, we can ask exactly how stable different categories have been.

8. Ronald Knox, introduction to *The Best Detective Stories of 1928–29*, ed. Ronald Knox, rpt. in *Murder for Pleasure: The Life and Times of the Detective Story*, ed. Howard Haycraft (New York: Biblio and Tannen, 1976), 251.

9. Gary K. Wolfe, *Evaporating Genres: Essays on Fantastic Literature* (Middletown, CT: Wesleyan University Press, 2011), 21.

10. For full metadata, see an online repository: Ted Underwood, “Data and Code to Support *Distant Horizons*,” Zenodo, last modified March 25, 2018, <http://doi.org/10.5281/zenodo.1206317>.

The story that actually emerged from this experiment didn't line up very neatly with either of the patterns we might have expected to find: generational succession or gradual consolidation. I see little evidence of the generational waves that Moretti's theory would predict. In fact, it isn't even true that books in a chronologically focused genre (like "the sensation novel, 1860–1880") necessarily resemble each other more closely than books spread out across a long timeline. Detective fiction and science fiction display a textual coherence that is at least as strong as the shorter-lived genres Moretti discusses, and they sustain it over very long periods (160 or perhaps even 200 years). So I think we can set aside the conjecture that twenty-five-year generational cycles have special importance for the study of genre.

But I also haven't found much evidence for the story of gradual consolidation that I expected to find. Although it is clearly true that the publishing institutions governing genre developed gradually, it appears I was wrong to expect that the textual differences between genres would develop in the same gradual way. In the case of detective fiction, for instance, the textual differences that distinguish twentieth-century stories of detection from other genres can be traced back very clearly as far as "The Murders in the Rue Morgue"—and not much farther. Detective fiction did spread gradually, in the sense that Poe and Eugène Vidocq were initially isolated figures, without a supporting cast of imitators, let alone genre-specific magazines and book clubs. But textual patterns don't have to develop as gradually as institutions do. Poe's stories already display many of the same features that distinguish twentieth-century crime fiction from other genres. Science fiction turns out to be almost equally stable, over an equally long span of time—which is a bit surprising, because the dominant critical story about science fiction strongly implies that it failed to consolidate until the twentieth century. These findings push in the opposite direction from much contemporary scholarship. Instead of being more volatile than communities of reception, textual patterns turn out to be, if anything, more durable.

Predictive Modeling

Computers enter this chapter largely to address a tangle of problems created by recent genre theory. If we could define genres once and for all by locating a single formal principle that unified them, the critic's task would be much simpler. We could say that science fiction is Darko Suvin's "literature of cognitive estrangement," and be done.¹¹ Unfortunately, readers rarely agree about the defining characteristic of a genre; different communities may choose to emphasize different features of the same works. Genre theorists increasingly suspect that genres are "family resemblances," constituted by a host of overlapping features.¹² Moreover, genres are historical constructions: the features that matter may change.¹³

In short, it increasingly seems that a genre is not a single object we can observe and describe. It may instead be a mutable set of relations between works that are linked in different ways and resemble each other to different degrees. A problem like this requires a methodology that is cautious about ontological assumptions and patient with details.

Predictive modeling fits the bill. Leo Breiman has emphasized that predictive models depart from familiar statistical methods (and I would add, from traditional critical procedures) by bracketing the quest to identify underlying factors that really cause and explain the phenomenon being studied.¹⁴ Where genre is concerned, this means that our goal is no longer to intuit a definition but to find a model that can reproduce the judgments made by particular historical observers. For instance, adjectives of size ("huge," "gigantic," but also "tiny") are among the most reliable textual clues that a book will be called science fiction. Few crit-

11. Darko Suvin, "On the Poetics of the Science Fiction Genre," *College English* 34, no. 3 (1972): 372.

12. Paul Kincaid, "On the Origins of Genre," *Extrapolation* 44, no. 4 (2003): 413–14.

13. John Rieder, "On Defining SF," 193.

14. Breiman, "Statistical Modeling."

ics would define science fiction as a meditation on size, but it turns out that the works librarians categorize as science fiction do spend a lot of time talking about the topic. Add clues from a few hundred more words, and you may have a statistical model that can identify other works librarians called “science fiction,” even if the underlying definition of the genre remains difficult to articulate (or never existed).

Hoyt Long and Richard Jean So have used predictive models in a similar way to recognize “latent, nonexplicit traces” of a haiku style in English poetry.¹⁵ The point of machine learning in projects like these is not primarily to enlarge the number of books we consider but to register and compare blurry family resemblances that might be difficult to define verbally without reductiveness. We can dispense with fixed definitions and base the study of genre only on the practices of historically situated actors—but still produce models of genre substantive enough to compare and contrast. Here, and in many similar cases, computation can help scholars acknowledge the complexity of culture without collapsing the conversation into an empty consensus that all cultural phenomena are more or less slippery and elusive. Which genres turn out to be especially mutable or relatively stable?

Since no causal power is ascribed to variables in a predictive model, the particular features we use to model genres are not all-important. The models in this chapter will take the frequencies of words and marks of punctuation as clues, along with a few other simple measurements (average word length, for instance). But we could use other features of the text if we preferred. The choice of features turns out to make less difference than readers encountering text classification for the first time tend to assume, because genre is expressed redundantly on many different levels. Critical debates about genre, for instance, may emphasize structural aspects of plot. But open a novel to a random page, and read

¹⁵ Hoyt Long and Richard Jean So, “Literary Pattern Recognition: Modernism between Close Reading and Machine Learning,” *Critical Inquiry* 42, no. 2 (2016): 266. For an earlier methodological model, see Jockers, *Macroanalysis*, 63–104.

a few paragraphs: without knowing the plot, you will quickly know whether you have opened a space opera or a detective story. Statistical models work by recognizing similar clues. Generally, efforts to improve genre classification by adding complexity to the model have turned out to be redundant: multiword phrases and information about character don’t significantly improve on the accuracy of simple lexical models.¹⁶

Moreover, if all the models in this chapter could be improved by 1%, it would make no difference to the argument. What matter are the *relative* strengths of the boundaries between different groups of texts. Which generic groupings are easier or harder to detect? Making our models as accurate as possible is worth doing only because it’s a reassuring sign that we haven’t left evidence on the table that would have changed those rankings. Of course, confidence of this kind can only be provisional. The methods I use to train models in this chapter (and throughout the book) are commonly used for text classification: regularized logistic regression on several thousand features, mostly the frequencies of words in the texts (see appendix B for details). I cannot prove that no better methods will ever exist. All I can say is that lexical models capture human judgments about genre rather well (accuracy above 90% is not uncommon), and researchers have been trying for decades to find a better strategy, without much success. As I explain in appendix B, I tried many different variations myself before settling on this strategy.

It is still conceivable that these models could have blind spots. For instance, we might expect lexical models to perform poorly with science fiction across a long timeline, since the specific innovations that count as science fictional have a limited life span. If we found many cases of this type, we would have to question the value of modeling genre with lexical evidence. But in fact, we find something close to the reverse. Human readers may disagree

¹⁶ Lena Hettinger, Martin Becker, Isabella Riger, Fotis Jannidis, and Andreas Hotho, “Genre Classification on German Novels,” 26th International Workshop on Database and Expert Systems Applications, 2015, <http://www.uni-weimar.de/medien/webis/events/tir-15/tir15-papers-final/Hettinger2015-tir-paper.pdf>.

about the continuity of Gernsback's "science fiction" with Verne's "scientific romance." But a lexical model has no difficulty grasping their similarity and is able to recognize science fiction across a long timeline with an accuracy of 90% or greater.

The challenges we encounter in modeling genre will usually come from excessive sensitivity, not from the coarseness we might expect to find in a quantitative model. The problem is not that lexical models fail to grasp the elusive similarities that define genres but that they turn out to be only too good at reproducing any category defined by human readers. Uncovering points of disagreement between readers will thus require a bit of ingenuity. For instance, if we fold the conflicting judgments of many readers together, an algorithm can usually still learn to recognize their areas of overlap and model the whole group of books with 70% or 80% accuracy. That won't necessarily prove that different readers agree. To get a more sensitive picture of the differences or similarities between readers, we will have to deliberately blind our model to some of the evidence and ask whether it can predict the judgments made by reader X using only evidence from reader Y.

Against Taxonomy

The word "genre" may evoke a mental image of a map that neatly partitions the landscape of fiction so that each work is located in one and only one region. Similar maps of musical genres have recently become popular online, and the rhetoric of supervised learning—with its references to "classification" and "accuracy"—may sound as though it begins from the assumption that cultural artifacts can be placed in a consistent taxonomy.¹⁷ In fact, our descriptions of fiction haven't created any such map, nor have I attempted to produce one here. A novel like *The Woman in White* (1859) is assigned by some observers to "the Gothic" and by others to "the sensation novel." In my metadata it bears tags associated

17. Glenn McDonald, "Every Noise at Once," accessed March 12, 2018, <http://everynoise.com/engenremap.html>.

with both categories. If I found it in a bibliography of detective fiction, I would add a tag for that claim too, connected to the particular bibliography making the claim. There is no limit to the number of overlapping genre claims that can be associated with a single work.

Other novels aren't associated with any determinate genre at all. In reality most nineteenth-century fiction was categorized very loosely, perhaps only by the phrase "A Novel" printed on the title page. If we needed to produce a symmetrical taxonomy of genres, we might redress this embarrassing situation by assigning uncategorized nineteenth-century works to a loose grouping like "realism." But why assume that there is any consistent taxonomy underlying human opinion? Let's begin with the inconsistent evidence before us. Instead of postulating that a taxonomy must organize the literary landscape, we can ask, empirically, how far human acts of categorization actually agree with each other.

This chapter is accordingly organized as a series of questions about the compatibility of groups defined by different observers. We will start by asking whether one definition of "detective fiction" has much in common with another. If we're able to get consensus on that point, we'll ask whether the category of "detective fiction" can be stretched further to encompass crime fiction or the Newgate novel. In each case we'll assemble a set of works tagged with a particular group of genre claims and ask how well it can be distinguished from a contrast set of equal size. Usually the contrast set will be selected randomly from a digital library (except inasmuch as it excludes the tags that defined the positive set). The contrast set will also be distributed across time in a way that matches the distribution of the positive set as closely as possible.

If there is no difference between the two categories being compared, one would expect the model to fail, making predictions that are no better than random (50% accurate, since the sets are always the same size). And indeed, if we attempt to model the boundary between two genuinely random groups of books, we get a reassuring failure, no better than random guessing. But few

categories produced by human selection are genuinely random. Even a very idiosyncratic or internally divided selection process will leave traces that can be modeled. For instance, if we mix works tagged as Gothic, detective, and science fiction, we produce a ghastly stew that few critics would call a coherent genre. But even this sprawling category can be distinguished from a random background, on average, 78% of the time. Glancing at a few of the most predictive features in the model—*murder, ghastly, lock, key, theory, and laboratory*—it is not hard to see how that happened. The genres involved here are not entirely dissimilar; a model can easily find sensational props or plot devices they have in common.

But genres occupy a space of similarity with more than three dimensions, so things that are very close along one axis can still be far apart in other ways. Although we can identify a mixture of detective, science fiction, and Gothic with 78% accuracy, the differences *between* these genres are even stronger than their collective difference from a randomly selected background. A model trained to distinguish detective fiction from the Gothic is right more than 93% of the time.

In other words, predictive models are rather like human beings: they can always find some ways two sets of works are similar and other ways they differ. If we want to know whether detective fiction and science fiction can be lumped together, no single model will answer the question. Instead, we need to pose a four-sided question that allows us to ask whether the differences defining one genre are parallel to the differences that define another. Do detective stories differ from other works of fiction *in the same way* that science fiction differs? To answer this question, one can train a model on the contrast between detective fiction and a randomly selected background and then ask the same model to distinguish works of science fiction from the same background. As we might expect, this model fails utterly: it's right less than half of the time. Although these two genres have a few things in common (theories and laboratories, for instance), their common elements seem not to be the features that most saliently define

them. When I need to decide whether two models of genre are similar, in the pages that follow, this is the test I will trust. Since I run it in both directions (asking a model trained only on A to recognize B, and vice versa), I call it “the method of mutual recognition.” This test certainly tells us that our ghastly “genre stew” can be separated out into detective fiction, science fiction, and the Gothic. The next question is whether “detective fiction” itself similarly breaks up into subgenres—groups of works linked by chronology or by a particular site of reception—that differ more than they resemble each other.

Detective Fiction

When researchers look for genres in their local library, they commonly rely on Library of Congress genre/form headings. These tags were applied to volumes by individual librarians and reflect tacit assumptions about genre held by many different people. They will provide an important source of evidence in this chapter. We can create an initial sample of detective fiction, for instance, by selecting works that librarians have tagged “Detective and mystery fiction.” But as we go back before 1940, these tags become very sparse, because we’re looking at books that were originally cataloged before the Library of Congress system assumed its present form. Only a few of those works have been recataloged in the last eighty years.

To create a sample of a genre from this earlier period we have to rely on bibliographies and critical studies. For pre-World War II detective fiction, I have relied on the catalog of an exhibition organized at Indiana University Library in 1973, covering “The First Hundred Years of Detective Fiction, 1841–1941.”¹⁸ This lists collections of short stories (and a few individual stories) along with novels. It also includes works in translation. I have been similarly inclusive throughout this chapter. Writers like

¹⁸. *The First Hundred Years of Detective Fiction, 1841–1941*, Lilly Library, Bloomington, IN, 1973, accessed 2015, <http://www.indiana.edu/~liblilly/etexts/detective/>.

Jules Verne and Émile Gaboriau strongly shaped the history of genre beyond France, so we would lose a lot by excluding them. I doubt, moreover, that there is anything untranslatable about the patterns detected by a model of genre: Verne will turn out to be an extremely typical figure within science fiction, even in translation.

It is of course possible that a single exhibition catalog of detective fiction (limited to volumes before 1941) will create a picture of the genre that diverges substantially from postwar volumes cataloged by many different hands. But that's exactly the kind of question statistical models allow us to test. Modeling just the eighty-eight volumes from the Indiana exhibition that I was able to obtain digitally, we have a rather high level of accuracy, 93.2%. The 194 volumes that have Library of Congress genre tags are more of a mixed bag and can only be recognized 91.0% of the time. If we combine both sets, we have 266 volumes (since 16 were in both groups) that can be recognized with 93.4% accuracy. So mixing groups selected in different ways doesn't reduce accuracy; it's a compromise that "levels upward."

But as I have mentioned, algorithmic models can be very good at finding common elements that link a group of works. A better way of comparing two human definitions of genre is to train a model on one reader's judgments and then ask it to predict judgments made by a different reader. For instance, when we ask a model trained on the Library of Congress detective fiction to distinguish the Indiana exhibition from a random background and vice versa, we still get 85% accuracy on average. That's the real confirmation that we're looking at largely congruent definitions of the detective story.

But what exactly *is* the definition of detective fiction operative in these models? It is not a shocking one. A quick glance at the words most predictive of detective fiction reveals the themes we would expect: *police, murder, investigation, and crime*. If we look a little deeper into the model, there are less obvious details. Architecture and domestic furnishing, for instance, loom large as a source of clues: *door, room, window, and desk* are all highly predictive words. On the opposite side of the model, words that

describe childhood and education (*born, grew, taught, children*) strongly predict that a volume is not detective fiction. Perhaps the genre's focus on a particular mysterious incident (or its tendency to take short-story form) encourage a contraction of biographical horizons.

It would be easy to put too much emphasis on these details. A model with four thousand variables is likely to contain many overlapping patterns, and even after we sort variables by their importance in the model, the important patterns will not all be contained in a few legible features at the top of the list. For instance, if we restrict our model to one hundred extremely common words, we can still identify detective stories with 86% accuracy, without any references to police or crime. The predictive words in this version of the model include vaguely interrogative signals appropriate for mysteries—*who, why, any, something*, and the question mark—but also more puzzling words like *have* and *was*. In short, the models created by machine learning are not mysterious black boxes: it is quite possible to crack them open and ask how they work. But they work so well, in part, because they are under no obligation to condense the world into a crisp definition. A model of a complex, overdetermined concept is likely to be just as overdetermined itself. Working with a light touch, a critic who has read a few detective stories can use predictive features to tease out insights (the absence of children in these stories, for instance, is an interesting dog that didn't bark). But it would be a mistake to imagine that a predictive model can quickly resolve arguments about the essential features of a genre. For that kind of condensing, prioritizing work, we still need critical debate.

On the other hand, statistical models are very good at comparing different frames of reference, and a method of that kind may be exactly what we need to address the skepticism about continuity now prevailing in genre theory. We have already seen, for instance, that a model trained on detective stories before 1941 is almost equally good at recognizing a collection of postwar detective fiction. How far can we stretch that sort of continuity?

Suppose we divide the history of detective fiction at 1915, before pulp magazines and writers like Agatha Christie or Raymond Chandler are thought to have given the genre its modern form. Can a model trained only on writers like Poe and Arthur Conan Doyle still recognize post-World War I detective fiction? Yes, without much difficulty. A model trained on examples published before 1915 is still 84% accurate with examples published on or after that date, and a model trained on the later set is right 87.8% of the time about the pre-World War I genre. This doesn't imply that there was no change across the timeline. We know that a great deal changed: the hardboiled style, for instance, hadn't yet begun to emerge in 1914. We could certainly train a model that would distinguish earlier and later examples of the genre. So the historian's standard chestnut—that continuity can coexist with change—holds true here, as it does everywhere. But our inquiry hasn't been posing a vague question open to that standard answer. We have been asking, more specifically, "Do the differences that distinguish detective stories from other works of fiction change enough to make twentieth-century examples of the genre hard to recognize?" And the answer to that is straightforwardly no. The things added to the genre, like hardboiled style, never efface the genre's other differentiating features. Scholars like Maurizio Ascoli, who argue that the contemporary genre has little to do with Poe, are simply not right.

The value of these methods is not just to affirm continuity. They are also sensitive enough to register discontinuity when it occurs. For instance, not every genealogy of detective fiction begins with Poe. Other influential arguments have traced the genre's concern with policing social order back to the Newgate novel of the 1830s or have found formal continuities with the sensation novels of the 1860s and 1870s.¹⁹ There is doubtless something to both claims, but the continuities involved are weaker than the

19. D. A. Miller, *The Novel and the Police* (Berkeley: University of California Press, 1988); Christopher Pittard, "From Sensation to the Strand," in *A Companion to Crime Fiction*, ed. Charles J. Rzepka and Lee Horsley (Oxford: Wiley-Blackwell, 2010), 105–16.

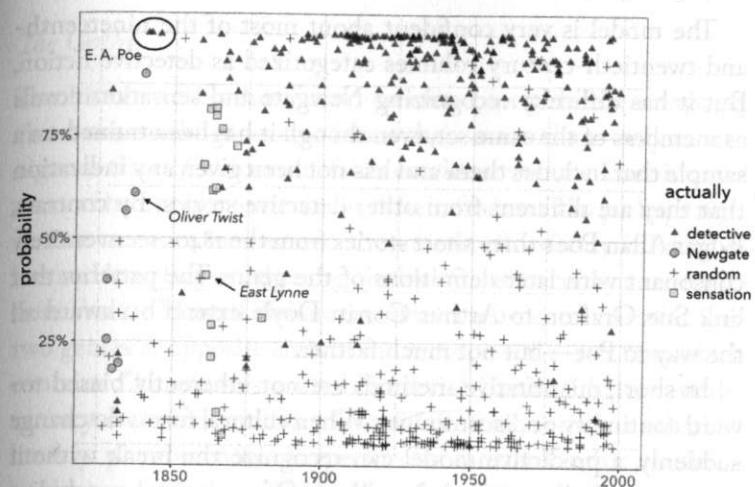


FIGURE 2.1. Probability of being detective fiction. Even when a model is trained on a definition of "detective fiction" that includes Newgate novels and sensation fiction, it tends to see those subgenres as outliers.

continuity with Poe. For instance, we can test these genres' similarity to other forms of detective fiction using the same method of mutual recognition we used to confirm continuity across the First and Second World Wars: in the case of the Newgate novel, this produces an accuracy of 59%, and for the sensation novel, 67.5%—distinctly lower than we found in earlier comparisons.²⁰ Alternatively, we can group all these genres into a single pool and ask which books turn out to be especially typical or especially hard to recognize. This question is easy to pose because statistical models naturally express their predictions as probabilities. So we can spread volumes out along a vertical axis that measures the model's degree of confidence that they belong to the positive set. In figure 2.1 I've done that with a model trained on 287 volumes from all the genres described above.

20. Keith Hollingsworth, *The Newgate Novel, 1830–1847: Bulwer, Ainsworth, Dickens and Thackeray* (Detroit: Wayne State University Press, 1963); Winifred Hughes, "The Sensation Novel," in *A Companion to the Victorian Novel*, ed. Patrick Brantlinger and William B. Thesing (Malden, MA: Blackwell, 2002), 260–78.

The model is very confident about most of the nineteenth- and twentieth-century volumes categorized as detective fiction. But it has difficulty recognizing Newgate and sensation novels as members of the same set, even though it has been trained on a sample that includes them and has not been given any indication that they are different from other detective stories. By contrast, Edgar Allan Poe's three short stories from the 1840s seem entirely consonant with later definitions of the genre. The patterns that link Sue Grafton to Arthur Conan Doyle extend backward all the way to Poe—but not much farther.

In short, quantitative methods are not inherently biased toward continuity or discontinuity. When cultural forms do change suddenly, a predictive model can recognize the break without being specifically instructed to do so. Of course, many scholars have already highlighted "The Murders in the Rue Morgue" as an important inflection point for crime fiction. Our model is not telling us anything radically new about Poe—in part, perhaps, because sudden innovations are things readers have always found easy to notice. But if we consider the other end of the timeline and compare the tight grouping of detective fiction from 1900 to 1949 to the more scattered pattern of detective fiction from 1950 to 1999, we may glimpse a more gradual change that has been harder to confirm on the scale of an individual reader's memory. It looks very much as if detective fiction is losing its crisp boundaries and starting to blur into other genres. This observation is made commonly enough about particular authors: Mercedes Lackey and Walter Mosley, for instance, both have a generic range that is difficult to pigeonhole. (And both authors do hover around the 50% mark in our model of detective fiction.) But when an argument is pinned to specific authors, it can be difficult to know whether the examples adduced are typical or exceptional. A model of 287 volumes puts us in a position to confirm that the similarities between works of detective fiction are really becoming looser, overall, in the 1990s than they were in the 1920s or 1930s.

The Gothic

The history of Gothic fiction creates, appropriately, a mystery about ancestral figures that haunt their descendants only as ambiguous traces. Critics seem fairly confident that the Gothic novel was a coherent phenomenon in Britain from 1760 to perhaps 1830. But as we move further into the nineteenth century, it becomes less and less clear whether the Gothic remains a continuous tradition. Franco Moretti divides the nineteenth-century Gothic into two genres at opposite ends of the century—the later one, which contains *Dracula*, designated "imperial Gothic." In twentieth-century America, "southern Gothic" is often treated as a distinct literary phenomenon. A specifically female Gothic tradition might run back through Du Maurier's *Rebecca* to Brontë's *Jane Eyre*. On the other hand, there are critical traditions that insist on the continuity of all these things and even stretch the Gothic to encompass the contemporary publishing category of "horror."

Since we have good reasons to wonder whether "the Gothic" writ large is a strongly unified tradition, it was particularly important in this case to compare different sources of testimony. Before 1840 I relied heavily on the Stanford Literary Lab's list of Gothic fiction; after 1840 I relied on the Library of Congress genre tags associated with "horror" or the "ghost story." But I also collected a set of works mentioned in *The Gothic*, a Blackwell guide edited by David Punter and Glennis Byron (2004), which traces a Gothic tradition all the way from Horace Walpole to Brett Easton Ellis, linked through a surprising range of intermediary figures that includes Emily Brontë, Henry James, and H. P. Lovecraft.

Few of these lists display the kind of coherence we found in detective fiction. The samples that can be predicted most accurately are the smallest: the twenty-one works (1791–1834) identified as Gothic by the Stanford Literary Lab can be recognized 86% of the time, and the twenty-seven volumes tagged "Ghost story" (1826–1920) can be recognized 91% of the time. The hardest

sample to model is the superset that combines them all: 171 volumes that can only be modeled with 81.0% accuracy. The growing shakiness of this Gothic edifice as samples grow larger betrays a weakness somewhere in its foundation. In modeling a homogenous group of works, accuracy ordinarily increases as one gathers more data. That is what happened, for instance, with detective fiction. Declining accuracy suggests that our different sources are describing heterogeneous things. We can confirm this using the method of mutual recognition. A model trained on Romantic-era Gothic works recognizes ghost stories only weakly and vice versa (69.8%); ghost stories are even less similar to twentieth-century horror (67.8%). If one compares twentieth-century horror directly to the Romantic-era Gothic, there is hardly any textual similarity at all (51.0%).

I don't expect this pattern to surprise many readers. Few critics claim that the Gothic is a genre as tightly knit as detective fiction. Even in the process of constructing a two-century anthology, Punter and Byron acknowledge the possibility that "there are very few actual literary texts which are 'Gothic'; that the Gothic is more do to with particular moments, tropes, repeated motifs that are found scattered . . . through modern western literary tradition."²¹ Establishing that the Gothic falls apart is a test less of the Gothic than of our method. A method that didn't recognize hard cases would be hard to trust about others!

I have characterized the Gothic as an ancestral memory, but the weakness of this genre is not simply a consequence of the passage of time. It is true that our timeline for the Gothic stretches back 240 years, 80 years further than for detective fiction. But making the timelines equal wouldn't render the genres equally coherent: there is no 160-year segment of the Gothic timeline that displays anything like the kind of coherence we saw in the detective genre. For that matter, there is little evidence that the 160-year history of detective fiction makes it more diffuse

²¹ David Punter and Glennis Byron, *The Gothic* (Malden, MA: Blackwell, 2004), xviii.

than shorter-lived trends. A group of detective novels randomly selected from the entire 160-year timeline can be modeled just as accurately as an equal-sized group selected from 25 years of the genre or an equal-sized group of sensation novels drawn only from the 1860s and 1870s. Generation-sized genres are not more coherent than others. We may instinctively assume that cultural phenomena should grow more diffuse as they cover longer spans of time, but the evidence doesn't always bear this out.

The fractures within the Gothic are caused not just by time but by the genre's reliance on several distinct thematic premises: ghost stories, Romantic-era tales of incarceration and escape (often mixed with a courtship plot), and twentieth-century horror. Each of these things is easier to model on its own than in combination with the others. Detective fiction is not internally divided in quite the same way. There are, to be sure, subgenres of detective fiction. In the twentieth century, for instance, I have used bibliographies to separate hard-boiled detective novels (Raymond Chandler, Mickey Spillane) from the sort of mysteries set in a country house with a closed circle of known suspects (Agatha Christie, Margery Allingham).²² A statistical model can easily distinguish these subgenres from each other: they are stylistically alien, in part because the hard-boiled style is histrionically gendered. Raymond Chandler openly defined the hard-boiled detective as a masculine (and American) response to the perceived femininity and intellectualism of British mysteries. But subgenres of detective fiction are never easier to model on their own than in combination. They can be at once very distinct (along a gendered stylistic axis) and very similar (along a different axis, which expresses a shared reliance on the criminal puzzle plot, in contrast to mainstream fiction).

In short, when we ask whether a genre is "unified," we are posing a complex question, which involves the depth and char-

²² John Charles, Joanna Morrison, and Candace Clark, *The Mystery Readers' Advisory: The Librarian's Clues to Murder and Mayhem* (Chicago: ALA, 2002); Geoffrey O'Brien, *Hardboiled America: Lurid Paperbacks and the Masters of Noir* (New York: Van Nostrand Reinhold, 1981).

acter of its internal divisions, as well as the extent of change over time. It is not safe to assume that all genres change at the same pace or that the existence of distinct subgenres will necessarily weaken a parent category. In the case of the Gothic, the forces of drift and division do work against coalescence. But even so, the term names a real phenomenon: a group of texts that can be recognized with 81% accuracy over a span of 240 years can hardly be called inchoate. Critics sometimes acknowledge the contrast with more tightly unified literary groups by calling the Gothic a mode rather than a genre. The term is fair enough, but like many terminological distinctions, it substitutes a crisp binary division for a more muddled reality. In the next section of this chapter, we will look at an intermediate case where the question of unity is not so easily resolved.

Science Fiction

I have argued that genre concepts that persist for more than a century can be just as coherent, linguistically, as those that persist for a few decades. But so far detective fiction is my only example, and there are reasons to suspect that the detective/mystery/crime genre might rely on an unusually stable set of premises. There's always a crime; there's always a detective; there's always an investigation. Science fiction would appear to pose a more challenging problem, because the premises of the genre are inherently mutable. The prototypes of the genre often describe conveyances like balloons and submarines that are no longer science fictional. Recent examples depend on information technology, which may produce a rather different plot. It is not immediately obvious that William Gibson's *Neuromancer* would share much common vocabulary with Verne's *Twenty Thousand Leagues under the Sea*. Many skeptical theories of genre have taken shape specifically around the mutability of science fiction—as the article “There Is No Such Thing as Science Fiction” reminds us.²³ In short, it

²³. Vint and Bould, “There Is No Such Thing as Science Fiction”; see also Kincaid, “On the Origins of Genre”; Rieder, “On Defining SF.”

is not intuitively clear whether we should expect science fiction to hold together over long timelines, like detective fiction, or fall apart like the Gothic.

Since there are different stories about the early history of science fiction, I drew on several different sources for that period. *The Anatomy of Wonder* is a well-known bibliography with chapters on early science fiction contributed by Brian Stableford, a writer of science fiction himself. Stableford's history of the genre strongly emphasizes H. G. Wells and the future-war tradition but is somewhat more reticent about other predecessor figures, like Mary Shelley and Jane Loudon. (Like many historians of science fiction, Stableford tends to define the genre through its scientific content, and he can be skeptical about works where that content seems lacking.) To get a fuller representation of women in the genre, I relied on a bibliography of women in early science fiction constructed by Mary Mark Ockerbloom. In spite of their different emphases, these sources construct lists of texts that can be modeled in very similar ways.²⁴

When all these bibliographic sources are folded together, we have a list of 213 volumes stretching from 1771 to 1999 that can be modeled with 90.6% accuracy. The boundaries of this genre are a little less clear than detective fiction, but it certainly has a coherence more akin to that genre than to the Gothic.

The scientific romances of the late nineteenth century are often described as groping toward a concept of genre that had not yet fully crystallized. But the model in figure 2.2 has no difficulty recognizing them as akin to Ursula Le Guin and cyberpunk. Three books by Jules Verne are all assigned to science fiction with a greater than 99% probability. Mary Shelley is a more ambiguous case. Her apocalyptic plague story, *The Last Man* (1826), is recognized confidently as science fiction, but *Frankenstein* (1818)

²⁴. Brian Stableford, “The Emergence of Science Fiction, 1516–1914,” and “Science Fiction between the Wars, 1918–1938,” in *Anatomy of Wonder: A Critical Guide to Science Fiction*, ed. Neil Barron, 5th ed. (Santa Barbara, CA: Libraries Unlimited, 2004), 3–22, 23–44; Mary Mark Ockerbloom, “Pre-1950 Utopias and Science Fiction by Women: An Annotated Reading List of Online Editions,” accessed 2015, http://digital.library.upenn.edu/women/_collections/utopias/utopias.html.

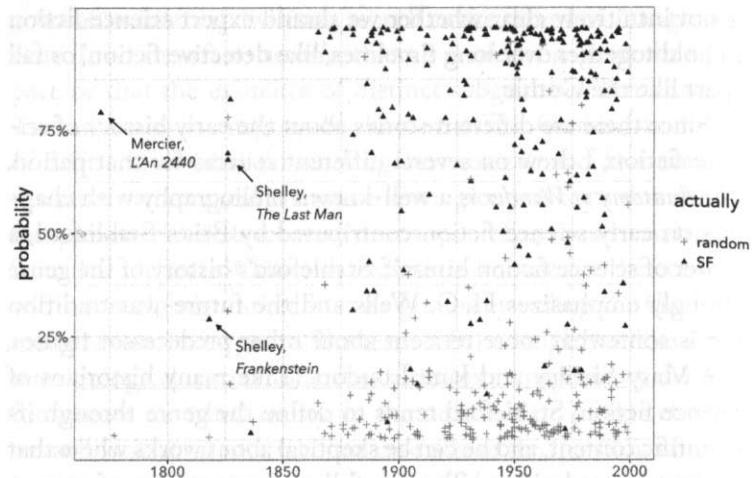


FIGURE 2.2. Probability of being science fiction. Science fiction, 1771–1999, classified with 90.6% accuracy.

hovers around the middle of the spectrum—as it does, frankly, in much critical commentary on the genre. So the boundaries of science fiction do become blurrier as we move back before Verne. Another site of ambiguity (not explored in figure 2.2) is the porous boundary between science fiction and fantasy. I have deferred that complex topic to a future article. Here it may suffice to say that science fiction and fantasy are separable traditions but traditions that have long been closely connected (more closely than either is connected to other genres).

Our list of science fiction, assembled by many different hands, includes cosmic evolutionary histories (Olaf Stapledon's *Last and First Men*) and personal dream visions (Marie Corelli's *A Romance of Two Worlds*); voyages into the future (Louis Mercier's *L'An 2440*), and into a prehistoric past (Arthur Conan Doyle's *The Lost World*). How is it possible to model such a long and varied tradition just by counting words? Science fiction turns out to have a strong stylistic signature, which we might loosely characterize as sublime. Invocations of scale (*vast, far, larger, infinite*) are very characteristic of the genre, as are large numbers

(*thousands, millions*). Horror, nightmare, and destruction are more prominent than one might have imagined. Self-conscious references to the *human* tend to accompany creatures against which *humanity* may be defined, and the pronoun *its* is common, since we often confront *unknown things* that lack an easily recognized human gender. At the other end of the scale, a whole range of quotidian details mark a book as probably not science fiction: references to *tea* or a *hat*, for instance, and to particular days of the week.

This is not by any means an exhaustive description of the model, and I want to be careful not to imply that sublime estrangement, by itself, is a sufficient definition of science fiction. This is just one salient and widely shared aspect of the genre's style. But this angle of analysis may help us understand why Mary Shelley's apocalyptic intensity belongs in our history of science fiction, with or without detailed scientific content. The same thing might be said about some of the randomly selected books that the model strongly (and persuasively) misclassifies as science fiction, like Thomas Pynchon's *Crying of Lot 49* and William S. Burroughs's *The Ticket That Exploded*. It may not be Pynchon's explicit concern with entropy but his paranoid fascination with the sheer scale of mass society that this model sees as connected to the tradition of science fiction.

One thing we might expect that doesn't appear in the model is the gradual consolidation of genre conventions that science fiction scholars spend so much time tracing. Historians of this genre are rarely as willing to give Verne as much credit as historians of detective fiction give Poe. The narrative premise of much historiography is that science fiction was an inchoate phenomenon (scattered across utopias, planetary romances, future-war narratives, and so on) until given a new shape and direction by particular American pulp magazines and anthologies between 1925 and 1950. Hugo Gernsback's *Amazing Stories* (1926) often plays a central role. Gary Wolfe says, for instance, that "science fiction, despite its healthy legacy throughout the nineteenth century, was essentially a *designed* genre after 1926." Even after that

point, “the science fiction novel persistently failed to cohere as a genre in the manner of mysteries and Westerns” until *The Pocket Book of Science Fiction* emerged in 1943.²⁵

As we will see in a moment, Wolfe and other scholars are probably right that science fiction changed rapidly in the middle of the twentieth century. But claims about failure to cohere are another matter: I don’t find much evidence to support them. The scientific romances of the late nineteenth and early twentieth century (1850–1925) seem to cohere perfectly well—in the sense that they resemble each other as closely and can be modeled almost as accurately (88.9%) as later examples of “science fiction” (90.2%). It is true that these ties of similarity are not quite as strong as the conventions organizing prewar detective fiction, but then, science fiction never becomes quite as predictable as detective fiction. I think one has to conclude that the early history of this genre looks incoherent to our retrospective gaze for parochial, presentist reasons. For instance, we are probably giving too much weight to the term *science fiction* itself, to the plausible scientific content that Gernsback demanded from his writers, and to a particular set of postwar technologies that we have come to view as iconic (robotics, spaceflight).

One could, admittedly, object that my list of nineteenth-century scientific romances comes mostly from twentieth-century histories of science fiction. Perhaps this group of texts is only coherent because it was selected, tacitly, to provide a set of precursors for a more strongly unified genre? This is a fair objection. In some cases this chapter has been able to found genre categories on nineteenth-century evidence: informal lists of Newgate novels and sensation novels, for instance, were constructed by reviewers at the time. But “scientific romance” was a looser category in the nineteenth century. Although the term was widely applied to Verne, Camille Flammarion, and H. G. Wells, I haven’t found a bibliography or reading list that I could use as strictly contemporary evidence for the boundaries of the genre. So there may

25. Wolfe, *Evaporating Genres*, 34, 21.

be something retrospective about the lists of nineteenth-century science fiction used here. But it is still significant that the concept coheres on a textual level. There is no way to construct the same kind of retrospective coherence for the Gothic.

If received stories about the incoherence of pre-pulp science fiction are misleading, is there, at least, some evidence that the pulps made a difference in the genre? Yes, there is, but measuring the pace of generic change is tricky. It might be tempting simply to measure the linguistic distance between science fiction in the 1920s and in the 1960s. But a simple measurement of distance risks conflating different questions. We wouldn’t want a broadly shared change in diction (say, the rise of contractions) to count as a change specific to science fiction. Instead, we need evidence that the distinguishing characteristics of the genre itself altered.

We can measure this sort of change by adapting the method of mutual recognition described earlier in this chapter. Sliding a window of sixty years down the length of a timeline, we can, first, model each thirty-year half of the window on its own, making generic predictions about works using contemporary evidence. Then we can switch the models: make predictions about 1930–59 using a model developed in 1900–29 and vice versa. Usually these predictions will be less accurate than the predictions that used immediately contemporary evidence. The amount of accuracy lost across any given midpoint (e.g., 1930 in our example) will give us one way to assess the pace of generic change across the whole sixty-year window.

In the case of science fiction, I think the story suggested in figure 2.3 matches received histories of the genre reasonably well. There are five lines for each genre because I ran the test five times, using a different randomly selected 85% of the books each time. For science fiction, the pace of change always peaks between 1930 and 1950, in a period that loosely aligns with Wolfe’s narrative of consolidation, bookended by the emergence of *Amazing Stories* in 1926 and the *Pocket Book of Science Fiction* in 1943. Detective fiction does not show the same midcentury acceleration of change—which is important, because it adds confidence that the

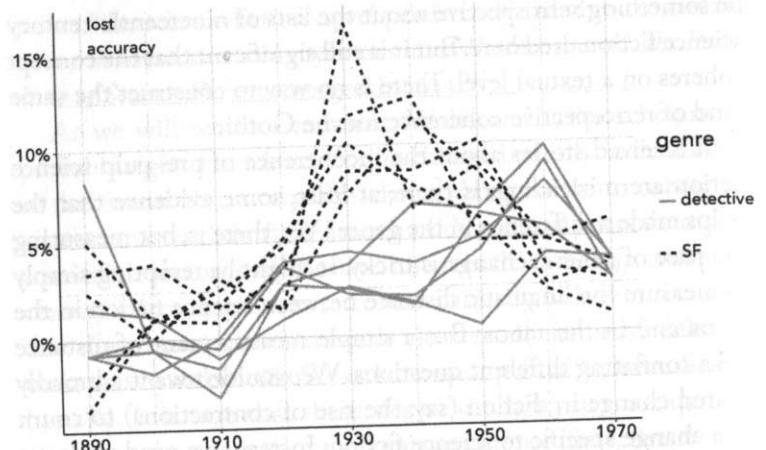


FIGURE 2.3. Pace of change in detective fiction and science fiction. Change is measured as the loss of accuracy when we use thirty years of the past to predict the future, and vice versa, compared to accuracy when each period makes predictions about itself.

shift is not produced by any purely mechanical change in our corpus (for instance, by changing the source of volumes from HathiTrust to the Chicago Text Lab at 1923). Moreover, I have recently replicated this experiment using a different set of volumes, drawn entirely from HathiTrust, and 1940 still turns up as the peak of change in science fiction.

So if science fiction did change rapidly between 1930 and 1950, what kind of change are we detecting? This doesn't have to be a mysterious question; there are many ways to open the so-called black box of machine learning and find answers. For instance, one way to locate change is to ask which later books are hard for models trained on an earlier period to recognize. Since models assign each book a probability of being science fiction, we can simply compare the probabilities assigned by models trained on different periods. This produces a measure of historical parallax or, so to speak, of surprise. If we sort science fiction in the period 1940–75 by this measure, clear patterns emerge. First, prewar models tend to be surprised by books of science fiction written

by women or with women as protagonists. Books by Catherine Lucile Moore and Ursula Le Guin are near the top of the list, along with Samuel Delany's *Babel-17* (with its protagonist Rydra Wong). But gender itself doesn't seem to be the full story: prewar models are just as surprised by Philip K. Dick and by Robert Heinlein's *Stranger in a Strange Land*.

If we wonder which particular aspects of these books stretched earlier conceptions of science fiction, we can pose the question by allowing different models to read brief passages. Then we can sort passages by the size of the difference between predictions made by prewar and postwar models. Many of the passages where the models disagree most sharply have a psychological or social dimension. The following example from *Babel-17* is typical. (I have italicized words that make prewar models particularly skeptical that this is science fiction.)

What "self"? There was no "I."

She had entered him in some bewildering reversed sexuality. Enclosing her, he was in agony. The light—you make! You make! his *crying* in terror.

Butcher, she asked, more familiar in patterning words about *emotional* turbulences than he, what does my mind in yours look like?

*Bright, bright moving, he howled, the analytical precision of Babel-17, crude as stone to articulate their melding, making so many patterns, re-forming them.*²⁶

There is a lot going on in this passage. Our model may not necessarily notice, for instance, that Delany has distorted English syntax to reflect his characters' struggle to communicate. The model does notice, however, that the central conflict of the passage is psychological: words like *emotional* and *crying* are legible clues, and they do a lot to convince models trained on prewar evidence that this passage is not science fiction. The passages of Le Guin's *Left Hand of Darkness* that are hardest for prewar models to recognize as science fiction similarly explore the psychological and social consequences of alien sexuality. And although Robert Heinlein's

26. Samuel R. Delany, *Babel-17* (London: Victor Gollancz, 1967), 159–60.

politics were different from Delany's and Le Guin's, *Stranger in a Strange Land* fails to fit prewar models of science fiction for the same reason: it pays less attention to the physics of spaceflight than to psychological, social (and specifically sexual) disorientation. Arthur C. Clarke, by contrast, is relatively easy for early-twentieth-century models to recognize as a science fiction writer.

Critics of quantitative approaches to culture often worry, with Martin Jay, that "there is no easy passage from micro- to macroanalysis."²⁷ That might be true, if macroanalysis meant simply counting words (as Jay assumes). A graph of macroscopic trends in word frequency can't tell us how the trend might have been produced by changes in individual books or paragraphs. But a predictive model of genre is another matter. Models are inherently relational, and one of their strengths is to build bridges between different scales of analysis—allowing us to understand how the historical contrast between two periods was expressed at the scale of the paragraph. I doubt the conclusions I have just drawn about *Babel-17* will surprise scholars of science fiction; the new centrality of psychological conflict in the 1960s loosely aligns with the well-known emergence of a New Wave in the genre. But this chapter's conclusions about long-term generic coherence and change are less obvious, and I felt it was important to show how those generalizations could be anchored in the recognizable stylistic quirks of individual paragraphs. Once we have a model of genre, that connection is not difficult to build.

In short, the so-called Golden Age of science fiction does seem to have been associated with an acceleration of change; works before that period are significantly different from the works published afterward. What isn't borne out is the common assumption that scientific romance before the pulp era was less coherent: a mere hodgepodge of literary premises struggling to coalesce around the genre concept that would eventually give them meaning. Nineteenth-century scientific romance may not

27. Martin Jay, "Hey! What's the Big Idea?": Ruminations on the Question of Scale in Intellectual History," *New Literary History* 48, no. 4 (2017): 626.

have had the institutional solidity or brand identity that science fiction later acquired. But our habit of equating brand identity with literary consolidation seems not to be well founded. The varied premises of late-nineteenth-century dream visions, voyages, and planetary romances don't prevent them from resembling each other, textually, as strongly as Golden Age works of science fiction. And although change did accelerate in the middle of the twentieth century, the changes in science fiction were never dramatic enough to prevent us from modeling two hundred years of the genre as a single unit.

Statistical Models and Genre Theory

The evidence gathered in this chapter challenges three existing theories of genre. Franco Moretti's conjecture that genre is a generational cycle is probably the least important of these targets: Moretti offered it as a reluctant speculation, and it has only been adopted by a few other scholars.²⁸ In any case, I have found no evidence to support it. The premise that genre boundaries gradually "consolidate" in the early twentieth century is a more serious matter. The notion that the pulps gave form to protean traditions that had previously "failed to cohere" is very influential in science fiction criticism.²⁹ But as we have seen above, the distinctive language of science fiction seems to take form before the institutions that are supposed to have consolidated it. The third theory of genre I have questioned is the recently popular notion that histories of genre are merely a genealogical thread linking a flux of disparate cultural forms. Predictive models can directly challenge this claim. If a model trained on detective fiction before 1914 can also recognize detective (and crime) fiction after that date, then then the differences separating the genre from the rest of the literary field must have remained relatively stable.

28. Adam Roberts, "A Brief Note on Moretti and Science Fiction," in *Reading Graphs, Maps, Trees: Critical Responses to Franco Moretti*, ed. Jonathan Goodwin and John Holbo (Anderson, SC: Parlor, 2011), 49–55.

29. Wolfe, *Evaporating Genres*, 21.

The historical picture that emerges from this inquiry is not simple. At first glance, there is a fairly clear contrast between long-lived twentieth-century genres (detective fiction and science fiction) and shorter-lived nineteenth-century phenomena (the Newgate novel or sensation fiction). But the rationale for this contrast is far from clear. Detective fiction and science fiction both seem to have stabilized before the modern marketing institutions that are usually offered to explain their stability. For that matter, it is not clear why certain genres stabilize more than others. Gothic novels and science fiction can both be traced back to the eighteenth century. It is not intuitively obvious why science fiction should have turned out to be the less mutable member of that pair.

In short, this chapter leaves many questions open. Its provisional conclusions could be revised by improving models or by adding perspectives from a wider range of observers. But even these tentative conclusions reveal that statistical models can challenge literary scholars' assumptions about the history of genre. Moreover, I hope it is beginning to become clear that the limitations of a quantitative approach lie in a different place than popular notions about mathematics might suggest.

Numbers are widely associated with a quest for objectivity in the physical sciences. In this story, the scientist is imagined to approach the world like Sherlock Holmes, "with an absolutely blank mind . . . simply . . . to observe and to draw inferences from our observations."³⁰ Humanists, for their part, have grown proud of unsettling that claim to objectivity. So any suggestion that numbers might illuminate human history immediately calls forth a well-rehearsed script, where math is expected to define objective patterns and humanists sigh that things are more complicated and depend on the observer's assumptions.

This script is rooted in years of conflict over the social role of natural science—in an era when numbers were good for measuring and counting but not for explaining.

³⁰ Arthur Conan Doyle, "The Adventure of the Cardboard Box," in *His Last Bow: A Reminiscence of Sherlock Holmes* (New York: Review of Reviews, 1917), 82.

suring physical things but almost useless for modeling human perspectives. Mathematical models that depend on social context may force us to reconsider the whole script. Numbers would do little to help someone who really approached the world, like Holmes, with an absolutely blank mind. If you want to count sheep, someone has to point at an animal first and say "that's a sheep." Similarly, to model genres I had to consult historically situated observers who pointed at Jules Verne and Samuel Delany, Agatha Christie and Raymond Chandler. Numbers did nothing to make this inquiry less perspectival. But they did allow us to compare perspectives and describe differences of degree between them. Once various observers have provided examples of the Gothic, for instance, we can use mathematical models to ask how compatible their assumptions are or how rapidly the genre mutates as we move down a timeline. In short, the point of quantification can be to render description relative rather than objective. Perspectival models are deeply compatible with an antifoundational approach to interpretation.