JUSTIFYING INTELLECTUAL PROPERTY

ROBERT P. MERGES

HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts
London, England
2011
I

ntellectual property (IP) law today is like one of those sprawling, chaotic megacities of the developing world—Mexico City, maybe, or Shanghai. Construction cranes are everywhere. The old city center—the ancient core of the field—is today surrounded by new buildings, new neighborhoods, knots of urban growth, budding in every direction, far off into the distance. As a longtime resident, an old-timer who for a good number of years now has walked the streets and taken in the scenes, I find myself with decidedly mixed feelings about all this. I marvel at the bold, new energy unleashed in the old burgh, and I am not a little pleased at the prosperity it has brought. But I also feel a distinct sense of unease. The helter-skelter of new growth, proliferating at times with no regard for the classic lines and feel of the old city, brings a slight case of vertigo—a feeling of being lost amid the familiar. It’s an exciting time, to be sure; but a confusing time too.

This book is a reconnaissance and renewal mission, undertaken in the precincts of my strangely familiar town. It is part archaeological dig, going back to the roots of the place to regain a sense of why it was founded, where it was centered, and what the city’s original outline looked like. It is also part mapping expedition, an effort to put in place a high-altitude conceptual map of the cityscape—one of those marvels of modern graphic design, with stylized lines for the main roads, and a color key for main features—something akin to the New York City subway map. The idea is...
Introduction: Main Themes

to cut through the noise and new construction, to locate the trunk roads and main boulevards that give this place its distinctive form and shape. And this book is, finally, an exercise in city planning: it lays out principles and guidelines drawn from the deep well of history, to structure and channel all the new growth. My goal is not to stop the growth, or to dictate all its features at a fine level. It is instead to make sure that with each new extension of the old city, basic themes and motifs from the historical core are picked up, replicated, and carried forward. As the city grows, I want it to retain its essential character.

The archaeological part of my mission is a search for foundations. Although all legal institutions emerge out of social practices and become formalized over time, an inquiry into historical origins forms only a small part of this book. The major archaeology I undertake is conceptual: in extending property to intangible items, what are the best justifications, and how do they shape the contours and limits of the field? In other words, what are the conceptual patterns, the basic formative ideas, that have inspired and animated the “cityscape” I am surveying? It is these questions, and not the fascinating (but for my purposes tangential) linear march of historical events (first this act, then this case, and so on), that motivate Part I of this book.

Current convention has it that IP law seeks to maximize the net social benefit of the practices it regulates. The traditional utilitarian formulation—the greatest good for the greatest number—is expressed here in terms of rewards. Society offers above-market rewards to creators of certain works that would not be created, or not created as soon or as well, in the absence of reward. The gains from this scheme, in the form of new works created, are weighed against social losses, typically in the form of the consumer welfare lost when embodiments of these works are sold at prices above the marginal cost of their production. IP policy, according to this model, is a matter of weighing these things out, of striking the right balance. At the conceptual level at least, the process involved is not particularly complex. It is easy to picture the toting up of costs and benefits, and to think of a good policy as one that equilibrates the scale at just the right point—the point that maximizes the number and quality of new creative works without costing society an arm and a leg.

The process is simple but, practically speaking, not at all easy. Impossibly complex, in fact. Estimating costs and benefits, modeling them over time, projecting what would happen under counterfactuals (such as how many novels or pop songs really would be written in the absence of copyright protection, and who would benefit from such a situation)—these are all overwhelmingly complicated tasks. And this complexity poses a major
problem for utilitarian theory. The sheer practical difficulty of measuring or approximating all the variables involved means that the utilitarian program will always be at best aspirational. Like designing a perfect socialist economy, the computational complexities of this philosophical project cast grave doubt on its fitness as a workable foundation for the field.

In my research, I have become convinced that with our current tools we will never identify the “optimal number” of patented, copyrighted, and trademarked works. Every time I play the archaeologist and go looking for the utilitarian footings of the field, I come up empty. Try as I might, I simply cannot justify our current IP system on the basis of verifiable data showing that people are better off with IP law than they would be without it. Maximizing utility, I have come to see, is not a serviceable first-order principle of the IP system. It is just not what IP is really all about at the deepest level.

This is a truth I avoided over the years, sometimes more subtly (for example, heavily weighing the inconclusive positive data, showing IP law is necessary and efficient, discounting inconclusive data on the other side), and sometimes less so (ignoring the data altogether, or pretending that more solid data were just around the corner). But try as I might, there was a truth I could never quite get around: the data are maddeningly inconclusive. In my opinion, they support a fairly solid case in favor of IP protection—but not a lock-solid, airtight case, a case we can confidently take to an unbiased jury of hardheaded social scientists.

And yet, through all the doubts over empirical proof, my faith in the necessity and importance of IP law has only grown. I seem to have a lot of company. Countless judges begin their IP decisions with one or another familiar “stage setter” about how IP protection exists to serve the public interest, often intoning one of a few stock passages penned in a spare moment by Thomas Jefferson. But these utilitarian platitudes quickly give way to doctrinal details, which often show the unmistakable imprint of something more fundamental, something beyond utility—revealing, at the end of the exercise, its real purpose and justification. That is, courts often wind up talking about IP rights as rights. Being courts, they are understandably too busy getting through their important work to notice the significance of this move. But in this book I make it my purpose to notice. And make no mistake, the shift from social utility to fundamental rights talk is an important one. For as we have learned from John Rawls, Jeremy Waldron, and others before them (particularly Immanuel Kant), the hallmark of a right is that social utility alone is not reason enough to override it. Waldron speaks to this when he distinguishes handily between “mere interests” and true rights. Despite the frequency of utilitarian rhetoric in the IP field, I have come to see that courts often understand IP rights (at
least implicitly) as rights in the full and true sense. Perhaps more surprisingly, at least to me, I have come to agree with them.

**Preliminaries: Is IP “Really” Property?**

I will argue in this book that there is a basic logic to the law of property, and that it applies to intangibles as well as physical things. Because there is a lot of resistance to this from various quarters, I had better state my best case here at the outset.

To begin, I must concede that on one narrow view of what “property” is, I cannot succeed. There are those who claim that property the concept is and always will be a prisoner of its origins, that it is rooted in and can never grow out of its formative association with physical, tangible things, most notably (in the Anglo-American tradition) land. Anyone who sees things this way will always see intellectual property as an awkward transplant. For them, the idea of property contains certain historical-essentialist traits that cannot be altered to better adapt it to intangible things. As applied to intangibles, property will always have the feel of a northern fir in the tropics, or a damp fern in the high desert. It just does not fit. And any conceptual adaptation, any strenuous breeding program that produces a concept that does fit would result not simply in a small variant. It would create something fundamentally new and different, a new species entirely; and whatever it was called, it could not be called property. (One lexical clue to this argument is that only property in land is called “real” property.) For those who subscribe to this theory, intellectual property can best be described as a sort of protracted analogy, a standing metaphor. It borrows some of the basic outline of property but cannot be considered seriously as an actual and true branch of this fundamental legal category.

I disagree. I do not see property in this historical-essentialist way. For me, property is a broad and roomy concept. It has a distinct (and fascinating) history, to be sure. But its origins do not imply constraints or limits. The very wide range of things that property concepts have been applied to suggests to me an expansive and highly adaptable legal category. Land, tools, trees, minerals, water, fractional ownership claims, legal obligations to pay money—these and many, many other things are subject to property’s wide embrace. Over its long career, property has shown a restless capacity to jump from one arena into another, morphing and adapting as it goes. While some of its distinctive features were shaped by its early history, I believe this history supplied property not with a set of burdensome constraints, but largely with a highly adaptable and flexible conceptual vocabulary that renders it wonderfully adaptive to all sorts of new things.
and situations. This vocabulary is singularly effective in structuring relations between legal actors and unique things of value to them. Property has proven robust because, like a spoken language that grows and spreads, it has shown itself quite capable of absorbing new dimensions and changing in significant ways, while retaining fidelity to certain core principles that provide its basic structure.

The most important core principle of the institution of private property is this: it assigns to individual people control over individual assets. It creates a one-to-one mapping between owners and assets. I argue in this book that this one-to-one mapping is the best way to handle intangible assets, just as it is with most other assets. For me, it is this powerful logic of individual control that makes property appropriate and appealing; it has little to do with the nature of the assets in question. That is why I see IP as a perfectly plausible, and even desirable, system for administering intangible assets. The logic of decentralized control and coordination—that is, individual ownership—makes just as much sense to me for intangible assets as it does for physical assets and the other objects of traditional property law.

Another major objection to the property model in IP law centers on the high transaction costs of property rights. Critics with this orientation do not dwell on essentialist arguments. Their point is more about the consequences of assigning property rights to intangibles, and in particular, the transactional bottlenecks and costly nightmares that this entails. There are many variations on this basic objection, but most share this thought: while IP rights may make sense for some things, or may have at one time, in the context of a rapidly evolving, high-throughput information economy, IP represents a major frictional “drag.”

I address this objection—with which I have a limited degree of sympathy—at several places in this book. My basic point throughout is that high transaction costs point to reforms in IP law, rather than the need to scrap it altogether. Compared to critics who emphasize this point, I am both more confident in the ability of economic players to “work around” transactional chokepoints and more convinced that, notwithstanding these transactional challenges, the best starting point remains a commitment to individual ownership of assets. Ultimately, I seek to learn from “transactional pessimists” but not join their movement.

Efficiency as One Midlevel Principle of IP Law

My earlier talk of IP as a “right” might sound like I am adopting a “natural law” perspective on IP, one that avoids completely any discussion of economic efficiency. Not so. Let me make clear right here that I am not
proposing that we banish questions of efficiency from IP law. Such a proposal would relegate a huge amount of helpful scholarship—some of it my own!—to the deleted file repository of history. The trick is to understand that efficiency is a second-order goal—a “midlevel principle.” Though I discuss foundational principles later in this introduction, I will take a moment now to explain the difference between a midlevel principle such as efficiency and a truly foundational concept.

Efficiency is an important goal of any area of law, and IP is no exception. The imprint of this important principle is all over IP law; indeed, many aspects of the social practice known as IP law cannot be effectively explained without reference to the principle of efficiency. As I mentioned earlier, however, despite its pervasive impact on the practices that make up IP law, efficiency is not an adequate foundational or normative principle. It cannot explain large features of the IP landscape (moral rights being one of many examples). And, try as we might, law and economics scholars have never established an efficiency-based (or utilitarian) justification for the field. There is no lock-solid proof that overall social welfare would decline if IP protection were suddenly removed. True, there are plenty of indications, plenty of data to support the notion that IP rights are overall a good thing for the economy. But there is no proof in the form that a scientist, or even a rigorous social scientist, would accept as unequivocal. The famous conclusion of the eminent economist Fritz Machlup in a study for the U.S. Senate was that it is not clear we would establish IP rights if starting from scratch today, but it would be unwise to get rid of them given that we already have them. The vast empirical literature in the field generated since then has done much to illuminate the wisdom of discrete practices and doctrines; but no one has produced evidence sufficient to dislodge Machlup’s basic conclusion. At a personal level, my interest (and belief) in efficiency led me to try to ground or justify the entire field on this idea. My failure to do so, and the path opened up by this failure, led me to write this book.

I also reject another principle that many have proposed (usually implicitly) as a possible foundation of the IP field: preserving and maximizing the public domain. Scholars have promoted this agenda under a number of rubrics. One centers on the idea that IP law serves the same function in the world of information that environmental law serves in the natural world: to guard as many things as possible from the rapacious grasp of privatization. In environmental law, this is the stewardship principle, the idea that our job is to protect nature’s wondrous but limited bounty from those who would appropriate it for personal gain. This concept lies behind environmental law’s preservationist agenda. In IP law, this concept is what I
call “the nonremoval principle,” which says that information and ideas in the public domain must not be taken away or privatized. This is the second midlevel principle that unites and helps organize the field.

Although I believe nonremoval is an important goal of the IP system, it is simply not robust enough to form a first-order principle. In this, it suffers from the same defect as utilitarianism. If you look carefully at various rules, doctrines, and institutional practices of the IP system, you will see much evidence of the nonremoval principle at work. And certainly for many practical purposes, promoting the agenda of nonremoval is a worthy aim. But just as with efficiency, nonremoval does not work as a normative foundation for the field. The reason is simple: it does not account for important practices and values evident throughout the field. For example, public domain concepts have nothing to tell us about the rules governing claims to priority (who was first, and deserves a right); the outcome of a priority contest in trademark and patent law is usually that one of the rival claimants will come away with a property right, so a policy favoring a maximal public domain has little place in such a contest. Likewise for the rules governing remedies and compensation when an IP right has been infringed. The core issue here is measuring harm to the rightholder, and nonremoval does not enter in. Even where nonremoval figures into a rule, it is often not the only principle at work. This is true with respect to rules about how much originality or creative spark is needed for a work to be worthy of copyright or patent protection. Nonremoval surely forms part of the rationale for these requirements. But another part comes from the idea that an IP right ought to be proportional to the contribution of a creative work. This idea, which I call “the proportionality principle,” is in my view central to IP law; it is perhaps the major midlevel principle and is covered in depth in Part II, Chapter 6, of this book. The fact that we need supplementary concepts such as proportionality is a dead giveaway that nonremoval is not by itself sufficiently robust to support the entire conceptual weight of the field.

It is plain to me now that although I thought I was starting at the bottom when I began exploring efficiency and nonremoval as foundations of the IP field, I was actually starting in the middle. In this I am not alone. The very accomplished legal philosopher Jules Coleman traces the same path in the introduction to his book *The Practice of Principle.* I use Coleman’s terminology (in Part II) when I describe the four midlevel concepts at the heart of IP law. Like Coleman, I believe that these midlevel principles serve a vital function. They tie together a whole range of disparate rules, doctrines, and institutional practices in the IP system. It is not surprising in retrospect that I would commence my review of IP law by looking into
these principles. They form an integral part of the connective tissue of the field. And in fact, as for Coleman, it was only by thoroughly grasping these principles and coming to terms with their limitations that I could begin to see the need for a foundational layer underneath them.

Before I introduce these foundations, however, I need first to complete my overview of the midlevel principles. In addition to (1) efficiency and (2) nonremoval, both discussed earlier, I have identified two more: (3) proportionality (mentioned briefly above), and (4) dignity.

Throughout IP law there is an impulse to tailor a creator’s property right in a way that reflects his contribution. This is the proportionality principle. There is a distinctly Lockean flavor to this principle (though it makes sense on utilitarian grounds as well), especially as Locke has been adapted to IP by legal scholar Justin Hughes. At its heart it is about basic fairness: the scope of a property right ought to be commensurate with the magnitude of the contribution underlying the right. Proportionality shows up in all sorts of IP rules, from infringement and remedies issues in copyright, to the requirements of patentability, to various trademark doctrines. It shows itself most clearly when a creator claims a right whose value is grossly disproportionate to the actual contribution at issue. In this situation, IP law finds a way to prevent the awarding of a disproportionate right. In copyright law, a very small piece of copyrighted material may hold the key to a large and lucrative market. “Lockout codes” on video game consoles serve as a good example. Ownership of a copyright in such a code can in effect translate into ownership of the right to control the market for games that are compatible with the console. Under some conditions, this exclusive control might make sense. But, according to several important cases, mere ownership of the copyrighted codes should not confer such control. The same principle is at work in patent rules denying effective control over large markets to creators who did not significantly contribute to the founding and growth of those markets. Opportunistic patentees often try to sail into a thriving market on clever strategies designed to capitalize on the pioneering work of others. But they usually run headlong into stiff countering winds. Sooner or later, judges and legislators push back against the unfairness of these strategies, on the premise that they are not consonant with the deeper purpose of the patent system. In Chapter 6 I use a simple parable and diagram—the Parable of the Bridge—to describe how the proportionality principle manifests in these cases.

The fourth midlevel principle in IP law is the dignity principle. It captures the idea that in many cases works covered by IP rights reflect and embody personal attributes of individual creators, therefore justifying special protection for some aspects of creative works. The dignity principle is
most evident in so-called moral rights, which protect the creator of a work even after other rights to the work are sold or transferred. The dignity interest can be thought of as an invisible string that connects individual creators with their works, and that survives even a formal act of legal alienation. For historical reasons, this principle is more fully developed in continental European IP systems. Even so, it finds expression (muted as it may be at times) in the IP law of the United States, where many (including me) feel it ought to have a more prominent place.

On Foundational Pluralism, or “Room at the Bottom”

Now that I have set out an overview of the midlevel concepts, I am eager to dig deeper and talk about what lies underneath, at the conceptual foundations of the field. But before I do, I need to digress for a moment. I want to say a few words about the relationship between foundations and midlevel principles.

Although I have arrived at my understanding of foundations over many years of study, I do not believe my ideas have any claim to exclusivity. The deontological foundations I describe in Part II are not the only plausible grounding for the field. As I said earlier, the current data (in my opinion anyway) are close to forming a lock-solid utilitarian case for IP. More data might tip the balance, leading me and perhaps others to believe that the field is basically all about net social utility, or perhaps that it can be justified by either set of core values, utilitarian or deontological rights. But this leads to a question: What would it do to our thinking about the field if the deep substratum could be changed under our feet—if in the face of new learning we suddenly substituted one “foundation” for another?

The answer at the operational level is: not much. That’s because the operational principles of the IP system are the midlevel principles I identified earlier. Efficiency, Nonremoval, Proportionality and Dignity—these basic principles, which form the conceptual backbone of the field, are largely independent of the deep conceptual justifications of IP protection. Except in a few boundary cases, they rarely do much direct work, or make a large practical difference, to the IP system in its day-to-day operation.

What about at the theoretical level? What would it do to our understanding of the field if we shifted from one foundation to another? I answer that by way of an analogy. In John Rawls’s book Political Liberalism, he does a masterful job of carefully pulling apart and then reconstructing the idea of pluralism. Rawls said that in a liberal democracy, there is need for a sort of
“public space” based on an “overlapping consensus” drawn from multiple, sometimes divergent foundational commitments. Through the proper construction of a liberal set of institutions, citizens can simultaneously hold to their deepest commitments (for example, fundamental values, religious faiths, and the like) and join together in a common understanding with others who hold equally strong but not necessarily congruent commitments. Liberal democracy, in other words, permits each individual to hold resolute personal beliefs while participating in civil society with others who may hold just as strongly to a contrary set of beliefs. And a really workable consensus, according to Rawls, includes not just a working agreement at the operational level but also a set of shared moral commitments, part of what Rawls terms “public reason.” These commitments constitute a level of “public moral discourse” separate from each individual’s deepest foundational commitments. According to Rawls, an overlapping consensus must include this level of moral agreement if it is to be robust—if it is to be flexible enough to adapt to new problems and situations.

My theory of IP includes this foundational pluralism. I am open to more or better evidence on the net social effects of IP protection. For me, a lock-solid utilitarian case might someday unseat deontological rights as the field’s foundation. In the meantime, the great virtue of pluralism is that I can engage in a meaningful way with those who are already convinced of the utilitarian account, those who hold firmly to deontological rights, and those who place their faith in other foundations altogether. Midlevel principles provide our common space, our place of engagement. They are like a musical score, allowing us all to play together, even if we disagree about the deep wellsprings or ultimate significance of our shared performance, our common musical practice. The midlevel principles allow us to be tolerant about questions of ultimate importance. In my theory, the conceptual hierarchy includes a ground floor that is airy and capacious. There is room at the bottom.

What about those who disavow the appropriateness of and need for any foundational theory whatsoever—those who would claim that the midlevel principles I discuss in this book represent the deepest theoretical level that IP law can and should aspire to? They might consider why the four midlevel principles I discuss (or any alternative set they might come up with) emerged as the appropriate organizing principles of the field. Isn’t it possible that a deeper metaprinciple informed the emergence and relative weighting of these principles? Understanding why a midlevel principle emerges in a field, and recognizing its limits and its relationship to other midlevel principles, can indicate the presence of a deeper organizing influence at work below the midlevel principles.
But a stout nonfoundationalist may of course reject this suggestion as well. Because many of my friends and fellow travelers in the law and economics community may feel this way, I want to direct a few words their way. Here is an analogy: Say you and I are scientists, old pals from way back who have worked long and hard in the lab in our chosen field. Suddenly I am plagued by doubts about where nature came from, why it is the way it is. I read widely and decide there is some sort of unseen intelligence, some higher power behind the whole thing, that set the universe in motion. My doubts go away and I rejoin you in the lab. My day-to-day work does not change, but I am somehow more serene about the whole enterprise, I have found a grounding for it that makes sense to me.

That’s what the normative grounding in Part I of this book is like—it helps me push on through foundational doubts. I am not saying that Kant and Locke are in any sense theological figures; just that they serve the same purpose for me as the spiritual-theological reading does for the scientist in my analogy. They provide a grounding outside the contours of my field as conventionally practiced, one that helps me resolve foundational doubts and get back to work confidently “inside” my field.

The important point is that I do not want you to think that this book, especially Part I, undermines my prior work or my commitment to analyzing detailed doctrines and rules, and the institutions that surround them, from the perspective of efficiency. In the vast majority of cases, the new normative grounding does not affect my view of correct policy in any way. It may help me resolve borderline cases, and perhaps might lead me to favor an owner or rightholder in a close case or at the margin. But mostly, it simply helps me frame the field. It gives me a stronger foundation for the conventional kinds of work you and I have always done. You may have no need for foundations. The system we have may need no deeper motivation or justification than that it has always had. My simple point is that this perspective was no longer sufficient for me. So I went on a deep exploration, the results of which are in Part I. Whether you see the need to join me or not, do not doubt that for the most part my views on right policy and the best way to look at our field operationally have not changed much at all.

Let us return now to the issue of foundations. I have already said that neither efficiency nor nonremoval is a foundational idea in IP law. But if not these, then what are the most fundamental ideas? First, IP is property. This may seem fairly obvious (it’s right there in the name of the field, right?), but given the current state of IP scholarship, it is not. IP scholarship is in the midst of a vast upheaval, caused in equal parts by the revolution in digital (and other) technologies and the rapid extension of IP law into these new
Introduction: Main Themes

technological domains: ownership of genes, Internet domain names, digital music “sampling,” open source software, and many others. Anyone who follows the popular press has seen the stories about the revolutionary challenges to IP law. This process of upheaval and extension has jarred many a scholar out of the view that IP is really property, or at least that property concepts are really still at the heart of the field. Many fine scholars (Larry Lessig, Mark Lemley, Tim Wu, and my colleague Peter Menell) question aspects of IP’s status as property; in so doing, they helped push me to think hard about the issue. They led me back to the field’s roots, while keeping me from a too-facile set of answers, the attractive (to me) but ultimately simplistic classical libertarian view of property associated with Richard Epstein (and Robert Nozick before him). While I would like to believe that a “night watchman state” is the ideal for IP, I see now that it is not. Lessig, Lemley, and the others have taught me well that optimal policy in the IP field involves more than simply providing a clear set of property rights and then getting the government out of the way. This is a good starting point, as far as I am concerned, but it is not the end of the story. Government must monitor more than the initial conditions of appropriation; it must keep track of how these property rights are assembled and deployed, and what consequences—economic and social—follow from their use in specific settings. Because property (like all rights) permits private individuals to bring the power of the state to bear against other citizens, the conditions surrounding the use of this power are always relevant, always of interest to the legal system. So the government’s interest in IP rights does not end when the rights are awarded. It extends to the long period during which property owners use rights to control and exclude other people. Acquisition and appropriation are only the first movement in the concerto of government concern with IP rights; they are not the finale. It is my position that treating IP as property is not in any way inconsistent with a concern for the environment and conditions in which property rights are deployed—the “postgrant” situation—and that this necessarily requires attention to distributional issues. Nothing in the label property limits me to a consideration only of the initial act of appropriation or grant, as some would have it. It is equally within the office of property as an institution to care about how much property is held by various individual people and legal entities, and how that property is used. None of this is in any way at odds with the simple idea that IP really is property.10

The second major realization that led to this book was that, if IP is really property, I should look for guidance in the literature on philosophical treatments of property. Books such as Jeremy Waldron’s The Right to Private Property,11 Stephen Munzer’s A Theory of Property,12 together with
Introduction: Main Themes

the books these works draw on—books by Locke, Kant, and Rawls—were all fair game. Although none of these works covered IP in any detail (most cover it very lightly, if at all), they all provided excellent grounding in the first principles of property. Collectively, they led me back to the principles on which this vital social institution is founded. More importantly, they led me away from a simplistic “just clearly define the rights and let the market sort things out” sort of view—the libertarian option I mentioned earlier. Instead, all these authors argue for a much richer, more complex, at times more confounding understanding of what property is all about. I came to see that they laid as much stress on the limits and exceptions to property claims (based most often on the needs of others, or “third-party claims”) as they did on the justifications for appropriating property in the first place. They led me back, then, to two connected ideas, which together form a truly liberal conception of property: first that property is essential to a fair society, and second that it is imbued with important limits and constraints. In writing this book I hope to translate these foundational writings into the IP context—to write a liberal theory of intellectual property law, with all that this means. This theory’s foundational components, as described by Locke, Kant, and others, are a commitment to individual ownership as a primary right, respect for third-party interests that conflict with this right, and, from the philosophy of John Rawls, an acceptance of redistributive policies intended to remedy the structural hardships caused by individual property rights. A liberal theory of IP rights will always remember that property is a grant from the government. It embodies an individual right, backed by social institutions—a right that is personal yet not selfish.

So my aim is to bring together social practices and institutions; midlevel principles, such as efficiency and proportionality; and the foundational concepts of Locke, Kant, and Rawls, in a single coherent theory of the IP field. Figure 1.1 summarizes the conceptual approach to the IP field that I describe throughout this book and serves as a high-altitude map of the terrain ahead.

So what are the first-order principles I am referring to? Stating them as succinctly as possible, they are (1) Lockean appropriation, (2) Kantian (liberal) individualism, and (3) Rawlsian attention to the distributive effects of property. I explain these principles in Part II, Chapters 2 and 3. For present purposes, I want to flag a few points from the later discussion that may help put these concepts into context.

John Locke’s appropriation theory is one of the traditional “first principles” of the field, unlike Kant’s autonomy theory. (The IP casebook, or
textbook, I coauthored begins, predictably enough, with an excerpt from Locke’s *Two Treatises of Government.* But I try to go a bit deeper with Locke than the main run of IP literature. In this introduction I can give only a brief sample of what I mean by this; the full account must be left to Chapter 2.

Starting with Locke’s famous “labor theory” of appropriation, I question the convention that he relies primarily on a metaphor of “mixing” one’s labor with materials found in nature. Applying labor to things, rather than mixing labor *with* things, seems a more apt description of what Locke was trying to say. This emphasis on applying labor achieves two important goals, one theoretical and the other more practical. At the level of theory, it relieves some of the difficulties, noted by critics, that are
caused by the emphasis on mixing. These difficulties are evident in Robert Nozick’s famous example of the person who claims ownership of the ocean after pouring a can of tomato juice into it. Talk of mixing shifts our attention away from the labor involved in mixing to the thing that is mixed; in Nozick’s example, from the effort of pouring the juice to the juice itself. The *reductio ad absurdum* flavor of Nozick’s example is the implicit comparison between the tiny volume of the juice and the immense volume of the ocean. How could adding the former justify a property claim to the latter? The absurdity disappears when we return to Locke’s intended subject, the expenditure of effort as a ground for property rights. He was not interested in trying to justify expansive property claims with scholastic treatments of tiny bits of new matter infecting massive amounts of old matter. He paid no attention to the relative weights or volumes of the “thing added” and the preexisting thing. He did not talk about what was added as a thing at all. He was concerned with *labor*—with how expending effort leads to justified claims of property. By making *effort* the subject of his inquiry, he naturally limited the scope of the resulting property claims. We need not get lost in the Nozickian metaphysics of admixtures, so long as our focus never strays from Locke’s primary topic: the labor itself. How much labor was involved? And how was the preexisting thing changed or affected by that labor? If we stick to these questions, we avoid absurd hypotheticals and keep to the spirit of Locke.

More practically, an emphasis on applying labor works much better for IP rights. The creator of a new work claims property not by virtue of contributing some new *thing* to a preexisting thing, but to the transformation of the preexisting thing by the expenditure of labor. Inventors, authors, composers, and the like do not add physical items like juice to the prior art or to culture; they add *effort* and, on that basis, claim property rights. A renewed emphasis on labor or effort therefore pays off handsomely when it comes to understanding the right to appropriate and its relationship to IP law.

After restoring Locke’s account of appropriation, I explore some other dimensions in Locke scholarship beginning with his treatment of third-party effects from appropriation—how Locke handles the interests of those impacted by appropriation. Contrary to the conventional view among IP scholars, I (following others, especially Jeremy Waldron) find Locke’s theory of appropriation to have crucial social and egalitarian strands. I believe this more egalitarian Locke to have important lessons to teach us about the proper reach of IP law, so I devote a good bit of space to these aspects of Locke’s thought in Chapter 2.

The basic foundations of IP law are individual autonomy and freedom. These are formative ideas from Immanuel Kant, so it might just as well be said that Kantian philosophy belongs, along with Locke, at the bottom of
the foundation. I must hasten to say that Kant was not much in my mind when I started this book. For me, reading Kant with an eye toward IP theory has been a revelation. I was aware of his foundational writings on epistemology, and—like everyone—I knew something about his seminal ethical theory (the categorical imperative, and all that). But reading Kant on property rights opened my eyes to a host of interesting ideas that I had only been aware of at the vaguest level. The relationship of property to individual will, individual choice, and individual freedom; the importance of possession as a concept, rather than an empirical fact; the idea that the formation of government (or “civil society”) precedes, rather than follows, the establishment of property rights (as in Locke); and above all, the marvelous vistas opened up by Kant’s characteristically conceptual approach to philosophy—these ideas washed over me as I was writing this book, slowly at first, and then insistently and rapidly. The real revelation for me was in discovering the excellent fit between Kant’s famous a priori concepts and the problems and contours of IP law. Having been schooled largely in the law and economics tradition of U.S. IP law, I was accustomed to thinking about IP law in the familiar terms of Anglo-American empiricism and utilitarianism. I was delighted and surprised to find a rich alternative. Kant’s conceptual approach—an approach based on carefully worked out deontological truths, rather than empirical facts and social practices—provided wonderful clues and insights to many of the questions that had long bothered and perplexed me about my field. Paradoxically, I also found that this conceptual approach is enormously helpful in understanding actual institutions and social practices—that the conceptual is a great aid to understanding the practical.15 I slowly realized why the conventional utilitarian account of IP law had been so dissatisfying to me. What I thought was the foundation of my chosen field of study, at least in its current state, could not bear the weight I had been asking of it. Something more fundamental was needed to support the structure of the field. Kant, with his emphasis on individual autonomy and the value of each person, provided just that foundation. And Kant’s style of thought is equally important. Looking back, it should have been apparent that there would be a good match between this most conceptual of philosophers and IP law, this most conceptual of legal fields (which Joseph Story, in the early eighteenth century, famously described as “the metaphysics of the law”). But it was not apparent, at least not to me. I hope as you read Chapter 3, on Kant and IP law, you will come to agree about the “goodness of fit” between Kant’s approach to philosophy and the foundational issues at the heart of the IP field. For now, a quick glimpse at some Kantian ideas will give the general idea.
Kant has complex ideas about creativity, ideas that track well with the structure of IP law. He begins with some primitive notions—the individual, his or her will, and the extension or application of that will onto objects. For Kant, the desire to shape and control things external to the self (that is, objects) is a powerful impulse for human beings. A project involving an external object may require that a person shape or control that object over a period of time. Therefore, human freedom depends, to some degree, on the ability to relate to an object in this way, to control and shape it over time. For some objects, this might be achieved by a persistent physical grasping, but this is obviously a limited strategy. Some objects are too big, hard to grasp, and so forth; generally, a more robust type of possession beyond physical grasping would be more effective in promoting the freedom to work on an object over time. Kant believes that this broader concept of possession is crucial to human freedom—so crucial, in fact, that it provides the impetus behind the creation of formal legal institutions, and hence civil society itself. For Kant, legal ownership is central to human freedom. Freedom, ownership, formal law, and then civil society: this is the key conceptual progression in Kant’s legal and political philosophy.

Contemporary theorizing about IP rights begins a long, long way from Kant’s system of thought, which is exactly why exposure to Kant can be so useful. Scholars today do not see individual freedom and the individual ownership it demands as the chief purpose of IP law. For most of them, IP law is strictly instrumental, a means to the ultimate end of net social welfare or the like. Kant cuts through this instrumental view as if wielding a knife blade. His thought upends amorphous concepts of collective interest and utilitarian balancing, replacing them with the bright, sharp idea of personal autonomy. The result is a more clear-headed focus on IP as a right, and on third-party interests as aspects or dimensions that are reached when we move outward from the starting point of the individual. Kant’s thought very effectively separates third-party interests from individual rights, a distinction I believe is essential to a proper understanding of IP law, especially at this point in the development of the field. An infusion of Kant promises to help correct the recent and intense emphasis on the rights of users and consumers of IP—a point I press in Part III.

Recasting IP in terms of Kantian rights does more than rebalance the field at the conceptual level, however. It leads to some immediate policy payoffs. Concern for autonomy, to take perhaps the most important example, goes beyond placing the rights of creators at the top of the legal hierarchy. It also means a thoroughly practical concern with the working conditions and economic prospects of creative professionals. Though this topic must await Chapter 7 for full development, the groundwork is laid.
in the discussion of Kantian property in Chapter 3. Autonomy is about something more than properly locating a set of legal rights at the apex of a conceptual pyramid. To be meaningful, it must have some cash value, so to speak; it must translate into putting a few dollars in one's pockets. Creative people are rarely free to create, and cannot effectively shape their destiny, if they cannot control and have little prospect of being paid for their creative work. Autonomy, it must be recalled, means "self-rule," the ability to steer oneself according to one's own plan and design. There is little chance of doing this in a sustained way without ownership over the products of one's creativity. Ownership confers both control and the prospect of compensation—the two practical dimensions of the abstract Kantian notion of autonomy.

So autonomy, in the Kantian sense, follows from conceiving of property as a right. But to speak of IP, or property more generally, in these terms (as a right) may sound like the beginning of a libertarian harangue. You may be familiar with the genre: an author makes a case for property as a right coequal with other fundamental rights; the sad plight of property, relative to other rights, is described (often with a disquisition on how modern governments have lost the way so painstakingly laid down by the early disciples of the "classical liberal" scripture); and the culprit (modern, redistributive government—the demon "welfare state") is identified, censured, and sentenced to a radical near-starvation diet, the better to return to its originally intended form, a sort of bantamweight overseer of robustly defined but minimally supervised entitlements.

Whatever else this kind of program is, it is not in my view consistent with Kantian foundations. For while Kant saw property as a right, and while all rights that deserved the title were fundamental, governments were absolutely not limited in purpose to defining or recognizing initial entitlements and then getting out of the way. Kant's Universal Principle of Justice pushed in quite the opposite direction. This principle, which serves as the baseline definition for justice in Kant's legal philosophy, says that actions (including claims to property) are fair or just only when "the freedom of the will of each . . . coexist[s] together with the freedom of everyone in accordance with a universal law." And there is no easy way to dodge concern for "the freedom of everyone" by arguing for the sufficiency of some sort of theoretical equal freedom to claim property, as against actual prospects for property acquisition and actual distributions of property holdings. Kant clearly rejects the easy libertarian path, calling instead for deep and true engagement with the interests of third parties. There is no getting around the difficulty of what Kant asks of civil society: to both respect individual property claims at a deep level and to simultaneously care
about the practical impact of individual property claims on the lives and fortunes of others. Kant’s thoroughgoing egalitarianism requires that legal rules relating to property maximize the freedom of all members of civil society.

This is a stringent pair of demands. Taken together, they seem to state something close to a paradox. It would take some real work to sort out, in a detailed way, how to reconcile these competing demands to arrive at a coherent set of principles for a property regime. Unfortunately, Kant did not see fit to do so. Outside of a few scattered examples of specific property issues (including an intriguing but brief passage on literary property rights), Kant’s work is silent on the details of how to balance the rights of appropriators and the interests of third parties.

Fortunately, just such a body of thought is ready to hand. It is the property theory of John Locke, and in particular the Lockean provisos and other limitations on appropriation. Kant’s more schematic ideas on meliorating the third-party effects of individual property rights lead straight back to the Lockean principles first discussed in Chapter 2. So in this way Kant’s work fits nicely with Locke’s, and Chapters 2 and 3 blend together to form a consistent description of the normative foundations of IP law.

That description is far from comprehensive, however. In particular, justifying individual acts of appropriation, and explaining principled limits on those acts, falls short of a comprehensive consideration of the overall societal impact of the institution of IP rights. For this comprehensive, systemwide view I needed the twentieth-century philosophical literature on property and distributive justice, and hence, of course, the work of John Rawls. In Chapter 4 I apply Rawlsian conceptions of the fair society to the problem of IP protection. The major question is whether individual awards of IP can be justified in light of their effect on the overall distribution of resources in society. A key question is whether it is fair to privilege creative works, which result from individual talent and effort, given that (according to Rawls) no one can be said to deserve their native talent, and also that societal resources are usually required to develop and apply that talent.

In addressing these issues I draw on the work of scholars who argue persuasively that dedicated development and application of talent gives rise to a legitimate desert claim, and hence, in my view, a justified claim to IP rights. Yet I also acknowledge Rawls’s original point, that much individual action is the result of pervasive social influence, so that society too has a legitimate interest—but not a coequal right—in the results of individual initiative. The societal interest is made manifest at three different points in the life of an IP right. First, and most conventionally, third-party interests are “baked into” the right; they form part of the structure of the
right from the moment it is granted. Time limits and affirmative user rights are examples. Second, the state rightfully monitors each IP right after its grant. Courts in particular intercede when postgrant conditions coalesce in a way that confers disproportionate leverage on a rightholder. User defenses and limits on legal remedies give primary expression to this impulse, which is described in depth in Chapter 6, The Proportionality Principle.

The third moment when societal interests enter into the IP picture occurs after a creative work has been exploited and money has been earned. In keeping with Rawlsian theory, society retains the right to claim some of the money that flows from individual creative works. Practically speaking, this means that it is fully legitimate for the state to tax the proceeds of individual works covered by IP rights. In addition to the Kantian core, reflecting the unique contribution of each creative person and the state’s recognition of individual autonomy, there is a Rawlsian dimension as well. I illustrate this balanced conception with a diagram showing the “deserving core” that belongs by right to all creators, and the “social periphery” representing society’s interest in each creative work.

This conception of individual property coupled with state-backed limitations and taxation is hardly novel. In fact, it combines two of the most conventional building blocks of Western socioeconomic systems. I emphasize this approach only because I think IP theorists often overlook it. Most choose to channel all their discussion of the proper balance between individual and society into the arena of IP law itself, as though each doctrine and each controversy must be engineered so as to get the balance right. It ought to come as a relief to be reminded that the conventional instrument of taxation is available to redress distributive imbalances, and that therefore we do not need to design every doctrine of IP law as a precision instrument designed to optimize overall distributive justice. The more systemic view supplied by Rawls’s way of thinking can get us out of the unproductive and often divisive trap of thinking that each individual rule of IP must balance out perfectly. Rawls’s approach frees us from this excessively internalist perspective and ought to be embraced for that reason alone.

I began by tracing the skeletal components of the contemporary IP cityscape, the midlevel principles. Next, I dug up the submerged foundations of IP law: the field’s formative concepts, derived from Kant and Locke, plus Rawls (with the caveat that other foundations are conceivable as well). Now I proceed to the material covered in Part III, a survey of some of the detailed institutions and rules that make up the surface and texture of IP law. I have two aims. First I describe some broad structural concerns that arise when discrete property rights, granted to individuals
for the foundational reasons I explain in Part I, and enforced and applied by courts according to principles laid out in Part II, come to be amalgamated and centralized in the hands of large corporate owners. Next I look at two important flashpoints in the contemporary discourse over IP law: digital works and pharmaceutical patents.

After this assessment of the way things are, I also look ahead in Part III. In Chapter 10, I draw on important critiques of IP—some substantive, relating to new technologies, others purely conceptual—that have appeared in the past fifteen years or so. Against some aspects of these critiques, I defend the field, arguing that IP rights continue to make sense even in the context of new technologies and conceptual challenges. At the same time, some aspects of these critiques are helpful adjuncts to the property-based traditions of the field. These I promote as useful wrinkles, new additions that cohere with and supplement the basic logic of property rights that has served society long and well.

The substantive critiques that IP experts have leveled in recent years touch on all sorts of issues. Two of the most frequent targets are copyright protection for digital works and pharmaceutical patents. I tackle these topics in Part III, where I extend and apply the concepts from Part I and the principles from Part II. I aim to show that these concepts and principles are more than useless artifacts or sterile taxonomic labels; instead, they are capable of bearing fruit, of doing real work in contemporary policy debates.

However, before we can apply the foundational concepts of Parts I and II, we must address a vital preliminary issue: what might be called the problem of corporatization. The principles and concepts we address early in this book are almost always described in terms of a stylized scenario involving a lone creator figure and his audience. Critics of these principles have often pointed out that the real-world conditions in the industries that “produce” IP-protected creations are far different. Very large teams of people work together, each member contributing only a small part of the larger creation, which is owned in many cases by a large corporate entity, often in conjunction with a large portfolio of other, similar works. The paradigm cases might be the Walt Disney Company for copyrighted films, and a large pharmaceutical or semiconductor company for patented inventions.

The reality of corporate ownership—and its sharp contrast with the idealized world of a lone creator—has a central place in the literature critical of IP rights. Indeed, one prominent critique argues that the romanticized notion of the heroic, lone creator was a myth created in the latter days of the Enlightenment and is perpetuated today by large companies that use it cynically to advance their own self-interest. Whether the strong form of this argument is true, there is a fundamental disjuncture between the individual
creator, who forms the stylized backdrop for foundational IP principles of the Lockean and Kantian variety, and the reality of large corporate “IP factories.” And for some critics, this disjunction is fatal; for them it means those foundational principles have little salience for IP policy.

In Chapter 7, Creative Professionals, Corporate Ownership, and Transaction Costs, I traverse and then reject these objections. My argument follows two main paths. I first point out that stories of the demise of the small creator are overblown; individual creators continue to make important contributions to culture and commerce. More importantly, if we widen the angle of vision just a bit, we can see that individual creators working in small teams—often in small start-up companies—play a very important part in contemporary IP-driven creativity. My initial thrust, in other words, amounts to a refutation of the basic premise of the critics’ notion that creative activities are now highly concentrated in large corporate entities.

The second part of my answer concedes that large corporations account for the preponderance of IP ownership in some industries. The question then becomes whether the foundational principles behind IP protection are still persuasive when large companies hold IP rights. Can IP law, based on rewarding individual effort and encouraging individual autonomy, retain its justification when IP ownership flows to a large corporate entity? The answer I give is a qualified yes. I begin by recognizing that corporate employers are not simply faithful agents of individual creators, and therefore that large IP-owning companies are not ideal “creator collectives.” There is, as a consequence, some dilution in the justification of IP rights. Put another way, it is important for the legal system to recognize that corporate IP owners mediate between the creative interests of the professional creators they employ and the business objectives of their managers. In some cases, this means that if IP law is to stay true to its foundational principles, legal rules must be tilted in favor of individual creators, at the expense of corporate owners. I give two examples. First, there are the rules that govern long-lasting copyright licenses, typically from small or independent creators to large corporate entities. A novelist who licenses “film rights,” for example, may argue later that the license should not be interpreted to cover an interactive DVD version of the movie based on the novel. Courts have struggled with these cases, at times interpreting the original licenses broadly and thus favoring the original licensee, and at times ruling the other way, in favor of the novel’s copyright owner. I come down in favor of proposals to systematically favor the individual creator in the interpretation of these licenses—the novelist in my example. A rule interpreting license grants narrowly—and therefore systematically favoring the small creator in this situation by giving him a chance to negotiate a
new, separate license for a new, unforeseen technology—comports best with the foundational principles of IP law.

The second legal issue that bears on the relationship between small creators and large corporate owners relates to the “rules of exit.” These are legal rules and doctrines that control how easy it is for an employee to leave a large company and start a new business. Because IP assets will often form the cornerstone of a new business start-up, ex-employers sometimes claim ownership of the ideas and technology upon which former employees found new start-up companies. Although I recognize the logic of large-scale corporate ownership, I also see the importance of liberal rules of employee “exit.” IP claims by large corporate entities should be carefully scrutinized to ensure they are not being used to prevent the formation of legitimate start-up companies. The misuse of IP along these lines cuts directly against the formative principles (effort, autonomy, and so forth) on which the IP system is based.

After acknowledging the problems of large corporate ownership, I consider and reject the stronger version of this argument: that creators are tiny peons in the larger corporate structure, and large corporate interests have so thoroughly co-opted the development of IP law that IP as it stands has become completely dissociated from the interests of individual creators. This is simply wrong. Though large corporate IP owners are not perfect agents of the creators they represent, the fact remains that large entities provide a professional home for many creative professionals today. Because of this, the interests of these large companies are often at least roughly aligned with those of the creative professionals they employ. This is why I advocate a more nuanced view than that of the full-time critic of large companies, one that recognizes there is sometimes a divergence between the interests of IP creators and the large companies that own much IP, but also that large organizations have an important role as employers and sometimes champions of creative professionals. I argue that corporate ownership is an important feature of the ecosystem that supports individual creative professionals; that incremental change in IP rules makes sense in some cases; but also that as long as new entry and a certain industry dynamism is possible there is no reason to suppose that industry structure is so anathema to the individual creator that the basic premise of IP law has become irrelevant.

These general thoughts about the structure of IP-based industries give way, in Chapters 8 and 9, to a discussion of problems and conditions in two particularly IP-intensive industries, digital entertainment and media (Chapter 8) and pharmaceuticals (Chapter 9). The idea is straightforward:
to apply the basic principles developed in the early parts of the book to some difficult contemporary issues—to put the theory through its paces, so to speak. In the process, I suggest some ways in which the foundational principles of the field can be brought to bear on some major challenges in these important industries. The idea throughout is to address contemporary problems with the tools and principles that form the basic building blocks of IP law, to confront new issues by remaining true to the property right foundations of the field.

In the context of digital copyright, the concepts of Part I and the principles of Part II yield two important insights. The first is a corrective, bringing our attention back from the places it has wandered to in contemporary discourse, which are all about the revolutionary nature of the Internet and digitization in general. After the basic grounding in the logic of property as a right granted to meritorious creators, it is easier to see the Internet for what it is: a revolution primarily in the means by which creative works reach their audiences. And this makes it harder to fall into a common trap, which is to get so caught up in this revolutionary distribution technology that we make it—and not the creators of digital works, the traditional beneficiaries of IP rights—the focus of our interest, and our policies.

It is as if society had developed a revolutionary new technology for car engines that makes it possible for the average car owner to go 500 miles per hour. A group of scholars emerges to argue that the goal of the legal system ought to be to enable as many people as possible to drive as fast as possible. Anything that interferes with this goal is “antitechnology.” Danger to pedestrians and bicyclists, concern for old neighborhoods and historic districts—these are irrelevant. The rhetoric of “regulation” and “control of technology” fills the pages of law reviews; the implicit premise is that any considerations that cut against the new technologies in any way are retrogressive, outdated, and reactionary. Arguing for rules that blunt the potential harm from this new high-speed capability brings immediate condemnation.

Just such a technocentric approach now dominates IP law, which is a shame. I argue that scholars should not get too caught up in the euphoria over the Internet. Instead, we need to keep faith with the basic ideas that have shaped the field, that give it structure. Scholars of the digital era have preached that “the Internet changes everything,” and indeed there have been huge changes in the way creative works are delivered to people, and in the things people can do with them once they have them. But the new technologies of dissemination have not changed one important fact: creative works still require effort, and (in many cases) a projection of the individual will, or personality. Since effort and individuality are the essence of IP,
property rights in creative works continue to make sense even in the Internet era.

Those who fail to see this essential continuity overlook the durable truths concerning creative works. A quick read of some books and articles in the technocentric vein proves the point. Lawrence Lessig’s *Code Version 2.0*,19 Jessica Litman’s *Digital Copyright*,20 and a host of similar writings repeatedly emphasize the radical discontinuity between the pre-Internet and Internet eras when it comes to technologies for distributing creative works. In this school of thought, the driving force of change is also the explanatory fulcrum on which the IP field hinges. The basic idea is simple: Because the technologies by which creative works are disseminated have changed drastically, our thinking about the field must change drastically as well. This is what I mean by technocentrism.

Technocentrism leads to two critiques of contemporary IP law. First is what might be called the unbalancing thesis. Scholars start with the idea that preexisting laws were formed on the basis of mostly unstated assumptions about what creators and consumers could do. Since those assumptions are now outdated—in particular, since creators now have much greater control over the uses of their works—the law has become unintentionally unbalanced. The idea (an old one in legal circles) is that the practical impact of a legal regime is a combination of “law on the books” together with real-world constraints and conditions, one of which is the state of technology surrounding application of the law. While the law on the books has not changed in many cases, these scholars argue that the practical impact or effective clout of the law has changed drastically. A common example concerns the distribution of written material—a magazine article, for instance. In the old, pre-Internet days, people could easily pick up a copy of a magazine at a bookstore or newsstand and read it for a while without paying for it. Libraries existed for the most part to facilitate shared access to individual printed works. Today, the owner of a digital version of the magazine can—and this is a crucial point—in theory monitor and charge for every use of the magazine, every user who glances at it or reads it over.21 So the literal application of existing law preventing the making of copies means, many scholars say, that the owners of copyrights today have far more and far stronger property rights in effect than in the old days.

The premise behind the unbalancing thesis is correct: the real-world environment plays a crucial role in determining the actual impact of legal rules. But there is a problem as it is applied. Scholars promoting this idea have failed to push it all the way through. They overlook the fact that although the physical/technological environment has changed now that creative works are routinely digitized, the business reality has changed as
well. So although the digital era has made possible much tighter technical restrictions on creative works, it has also made it much more important for copyright owners to permit all sorts of free sampling and reuse of creative works. This sort of liberality is dictated by the market: users want this sort of freedom, and owners must give it to them if they are to have any chance of seeding and developing the market for their works.

Chapter 9 considers the complex issue of patents on pharmaceuticals of use to suffering populations in the developing world. This is a classic conflict, pitting the rights of pharmaceutical companies against the ethical claims of a destitute group. The ideas developed in Parts I and II can be put to good use in sorting out the respective claims involved. Though Locke and Kant provide support for the basic notion that property makes an appropriate reward for the research effort of pharmaceutical companies, it is the limiting principles and constraints from their theories of property that really come to the fore. Locke’s charity proviso, for example, is directly relevant. Under this proviso one whose survival depends on access to resources controlled by another has a legitimate property right to the resources. Locke’s notion of charity thus extends to the point where destitute populations whose lives are truly at stake have a strong claim to life-saving pharmaceuticals. Likewise, Kant’s Universal Principle leads to the same result. The effects of pharmaceutical patents on destitute populations represent an extreme limitation on their autonomy, and so the rights of the patent owners must give way.

There is however a significant boundary around the access rights of the destitute. Severe and persistent inroads on the patent rights of pharmaceutical companies may threaten the long-term viability of the drug industry’s research program. Future generations may suffer the consequences. Under what Rawls called the “Principle of Fair Saving,” more generally known as the problem of intergenerational equity, it would be wrong to redistribute resources today so as to threaten the welfare of future generations. In the case of pharmaceutical patents, this implies the need for limits to the right of access that follows from a straightforward application of the Lockean and Kantian principles to this case.

Chapter 10 serves as a précis of the main arguments in the book. Its primary point is that property still makes sense. The one-to-one mapping of individual owner to discrete assets remains a crucial and compelling social institution even in the modern era. There remains a strong argument for recognizing and rewarding creative work with true legal rights, thereby converting that work from hourly wage labor into a freestanding economic asset wherever possible. Nurturing professional creators means en-
Courage not only individual and small-team ownership, but also large corporate entities, which form an important part of the ecosystem that nurtures and supports individual creative professionals.

At the same time, I reiterate in Chapter 10 that just because IP rights should be real rights does not mean they must be absolute rights. In addition to the built-in limits on appropriation suggested by Locke, Kant, and the midlevel principles of Chapter 5 (chiefly nonremoval and proportionality), there is society’s right to levy taxes on the proceeds from works covered by IP rights. As explained in Chapter 4, this is one of the most effective ways to maintain IP rights as an institution while recognizing society’s contribution to and interest in each IP-protected work.

Limiting property rights and taxing IP-covered works are not the only ways to accommodate the needs of consumers and users. As I explain in Chapters 7 and 8, I favor cheap and easy IP permission and licensing mechanisms, together with simple waiver techniques that permit binding dedication of rights to the public. In the long-standing debate over incentives versus access, creators/owners versus consumers/users, I argue that this is the right combination. Rightholders can continue to receive rights, while consumers and users can gain access to the works they want to use, if resources are directed to creating efficient transactional mechanisms that allow IP rights to flow through commercial channels as smoothly, or almost as smoothly, as do the works covered by those rights. We need to recognize that, in a world with numerous IP rights, the market for creative works necessitates also a (separate, but related) market for the rights covering those works, and that IP policy ought to encourage market making in this secondary market.

I also call for a simple and binding mechanism for waiver—allowing a rightholder to make a binding dedication of his works to the public and thus implementing a “right to include” that is coextensive with the traditional right to exclude at the heart of IP, and property generally.