

Introduction

Until fairly recently, the law of intellectual property – a term that encompasses patents, trade secrets, copyrights, and trademarks, among other things – was something of a backwater. Of interest mostly to specialists within these fields, it garnered little attention from the broader legal community. Most economists manifested a similar indifference to these issues. Only a few gave serious consideration to the design of patent rights and even fewer paid much attention to trademarks or copyrights. Indeed, some law-and-economics scholars doubted that economics had much to say about any of these bodies of law.¹ Roughly within the last ten years, due in large part to the expanding role of high technology in our everyday lives, all of that has changed. The law of intellectual property – particularly patents and copyrights, but also trademarks and trade secrets and other related fields – has become a topic of major interest to lawyers, judges, and law professors. Many high-profile cases are making their way through the courts; new legislation is being introduced in many countries and international treaties are attempting to properly balance the incentives for investment against the need for access to the products of that investment, such as essential medicines. Economists have also taken up the challenge of modeling the consequences of high and low levels of protection and, to some extent, of testing these models against the empirical evidence. It is now common for leading law reviews and economics journals to publish articles on these issues – to say nothing of the popular press, with its endless fascination for such items as the attempt to patent the human genome, the ongoing controversy over the digital distribution of sound recordings and other copyrighted works,

¹ For example, in 1986, Yale law professor George Priest wrote that “economists can tell lawyers ultimately very little about how to enforce or interpret the law of intellectual property.” See Priest (1986; 21).

and cases that push the envelope of trademark protection for remote source signifiers such as product design and color.

However, even within the burgeoning literature on the law and economics of intellectual property rights (IPRs), there is still relatively little discussion of the appropriate remedies for the infringement of patents, copyrights, trademarks, and trade secrets. There is, to be sure, a fairly widespread consensus that an injunction – an order to cease infringing – is the appropriate remedy in most cases in which the plaintiff proves that the defendant has trespassed the plaintiff's rights. But there is relatively little discussion of the law of damages, and this gap is curious. Even in a system that routinely grants injunctive relief, damages are a necessary remedy for the time period running from the beginning of the infringement to the entry of the injunction. And this may be quite a long time, depending on how difficult it is to detect infringement; the ease with which litigants may obtain preliminary injunctive relief; and the substantial time it takes for a case to go through the legal system. A system that awards very substantial damages in effect strengthens the owner's IPR, whereas a system that awards minimal or no damages, or that imposes insuperable difficulties to the proof of damages, necessarily weakens those rights. More generally, the remedies that are available for infringement in some respects drive the entire IPR system. Without effective remedies for the enforcement of these rights, the rights are worthless; on the other hand, if remedies and other enforcement mechanisms are too generous, they may cause the cost of protection to be raised to a point that outweighs the potential benefits. Inattention to remedies, in other words, can undermine the whole system, no matter how much careful thought and analysis have gone into devising the rules of substantive law.

WHERE THIS BOOK WILL TAKE US

This is the first book-length treatment of which we are aware of the law and economics of remedies and other closely-related issues in intellectual property (IP) law.² We begin in Chapter 2 with an overview of the law of patents, trade secrets, copyright, and trademarks, and of the economic rationales for (and critiques of) the principal features of these bodies of law. As this discussion will show, all of the various bodies of IP law ideally strike a balance between incentives (to create, to publish, to invest in product

² A just-published book by Landes and Posner presents a detailed economic analysis of many issues of IP law, but specifically disclaims a systematic analysis of the law of remedies. See Landes & Posner (2003; 7).

irce
nics
ion
hts,
on-
iate
has
the
nts
un-
on.
tect
nc-
egal
ens
, or
rily
for
fec-
ess;
too
hat
rds,
and

law
tual
law
mic
s of
ally
luct

many
See

quality), on the one hand, and public access to the work product that results from these incentives on the other. Just where that ideal balance lies remains a matter of disagreement. Scholars are divided on the issue of how strong IPRs should be, particularly in the digital environment and with respect to new technologies. But our goal in this book is not to resolve these issues. Instead, we will assume that the policymaker has chosen a particular scope and duration for the IPR at issue, and that this choice reflects some reasoned consideration about the proper balance of social benefits and costs. We then offer some insights as to the advantages and disadvantages of different possible rules for the private enforcement of these rights in court. As we will see, some of these enforcement rules may function better than others at preserving the incentive structure embedded in the substantive law; others may be less costly to apply, but may not function as effectively at preserving that structure.

More specifically, we want to be able to answer the following sorts of questions. First, with respect to remedies, as we noted earlier, IP law evidences a marked preference for injunctive relief. Rules relating to damages and other forms of monetary relief are nevertheless also a necessary supplement, if only for those cases in which injunctive relief cannot be obtained immediately (or at all). But what sort of monetary relief should courts provide in order to preserve the incentive structure upon which the system is premised while at the same time avoiding the overdeterrence of lawful conduct? Is a lost profits or lost royalty remedy sufficient or should courts in some cases award relief in the amount of the defendant's gain instead of in the amount of the plaintiff's loss? How can one measure the amount of the plaintiff's loss (or defendant's gain) that was attributable to the act of infringement and not to other factors? If neither form of monetary relief can be calculated with confidence, should courts opt for some form of fixed or presumed damages, as in defamation cases? Is there a role for punitive or other super-compensatory relief? We address questions of this nature in Chapters 3 and 4, and then return to some measurement problems in Chapter 8.

A second set of issues with respect to enforcement relates to the proof of infringing conduct. Even if the defendant's invention, work of authorship, or trademark falls within the scope of the plaintiff's rights, should the defendant's liability be conditioned upon proof of a particular mental state (such as intent or negligence)? Or should liability in these cases be strict? U.S. IP law is often referred to as a body of strict liability law, but as we will show what we really have is a *sui generis* system in which the defendant's ability and effort to discover the plaintiff's entitlement *ex ante* has some bearing on the relief to which the plaintiff is entitled. Although the rules vary from one

body of law to another, the remedies afforded for the infringement of IPRs are often conditioned upon proof of some sort of knowledge or notice or other mental state on the part of the defendant. This insight is trivially true with respect to injunctive relief. No court will enter a preliminary or permanent injunction, prospectively enjoining the defendant from infringing, unless and until the defendant has been served and has had an opportunity to be heard. But it is also, less obviously, often true with respect to damages remedies as well. Patent law, for example, conditions awards of damages upon proof of at least constructive notice (most of the time, at any rate) and copyright conditions all relief upon proof of copying (albeit sometimes unconscious copying is sufficient). Along the way, we discuss the merits and demerits of different possible liability regimes, including a “pure” strict liability regime, a negligence regime, and an intent-based regime.

A third set of issues relates to the identity of the proper parties to an infringement suit. U.S. patent law, for example, extends liability all the way down the chain of distribution by rendering manufacturers, sellers, and users jointly and severally liable, subject to the first-sale doctrine. Copyright and trademark law traditionally have limited liability to a greater extent, although some recent developments tend to move these bodies of law in the direction of the patent model. Is there any logic to the way in which these bodies of law have decided who should be responsible for the act of infringement and who should not? Similarly, is there any logic to the various ways in which IP law allocates the right to sue for infringement? Here, patent law is the most restrictive in allocating the right to sue; it essentially restricts suits to patent owners and their exclusive licensees who must join the owner as a party. Copyright law is the least restrictive and permits suits by any “beneficial owner” of a copyright interest, including owners and exclusive licensees of individual copyright rights falling short of the entire “bundle of sticks.” The trademark rules are less easy to describe in a sentence or two but might be viewed as falling somewhere in between the rigorous patent and less rigorous copyright rules. Are there good reasons for these differences? Should the rules be modified? As we shall see, the answers to these questions depend in part upon the various ways in which rights may be licensed and in part on the likelihood of a right being invalidated in court, which varies from one body of law to another.

We anticipate two possible critiques of our analytical framework. One critique is that our analysis, which takes the existing scope and duration of IPRs as a given, cannot improve social welfare because the existing scope and duration of IPRs are suboptimal. To cite just a few problems, the duration of copyright is probably much longer than it needs to be under the incentive model; the optimal breadth of patents and copyrights is uncertain;

?Rs
or
rue
er-
ng,
ity
ges
ges
ite)
nes
rits
rict

an
way
ind
ght
ent,
in
tch
t of
ous
ent
icts
ner
any
sive
e of
but
and
zes?
ons
and
ries

One
n of
ope
du-
the
ain;

and the one-size-fits-all nature of much of patent and copyright law may be unfortunate as well, because equal terms are conferred on both high- and low-value products. To the extent that the current system is suboptimal, our suggested rules for preserving the incentive structure embedded in that system will only exacerbate the problems or not go far enough toward correcting market failures. A second critique is that, even if our task is worthwhile, it is impossible, because matters of enforcement and procedure cannot be sharply contrasted with matters of substance. This is perhaps most evident in connection with our discussion of liability standards: if a particular mental state is required (or not) for an act to be deemed infringing, then the scope of the IPR owner's rights is narrower (or broader) than if the rule were otherwise. But the point can be made in connection with the other chapters as well. A rule that awards the patent owner her lost profit confers a broader scope than does one that awards only a fraction of that lost profit, because in the latter instance some acts of infringement may be profitable to the infringer, and therefore will occur even though illegal; in effect, patent scope has been diminished under the latter rule for good or ill.

We recognize the force of these critiques, but nevertheless adhere to our basic framework for two reasons. First, while our analysis assumes the optimality of a given system of IPRs, it does not depend upon any *particular* system being optimal. In other words, even if the copyright term was shorter, the patent term was longer, and the scope of both bodies of law was more precisely calibrated with the maximization of social benefits, one would still need to confront the issues we discuss in this book. Moreover, one would – we think – still reach the same basic conclusions. Therefore, our analysis is, to borrow a metaphor from computer technology, platform-independent. As long as IP law is viewed as embodying certain incentives designed to maximize social welfare, one will need to craft enforcement rules so as not to undermine those incentives. Our analysis suggests a variety of ways of doing this, even if the current mix of incentives can itself be improved upon.³

³ Another way of articulating our project would involve the positive/normative distinction that is commonly found in discussions of law and economics. In one sense, our analysis is largely positive, because it asks whether current enforcement and procedural rules are consistent with or help us to better understand the incentive structure that in theory is embedded in existing substantive law. Whether the existing substantive law is itself optimal is another matter. From another angle, our analysis might seem more normative, because it asks what the enforcement and procedural rules *should* look like, assuming the existence of a proper incentive structure is embedded in the substantive law. We think that the previous description of our project conveys what we intend to do and why we intend to do it, regardless of the label one chooses.

Our second response is jurisprudential. The current system may be sub-optimal and by various maneuvers one could manipulate the rules relating to enforcement to better align costs and benefits. For example, if one thinks that the patent system in its current state confers too many costs and too few benefits, one might try to remedy the situation in part by making it extremely difficult for patent owners to recover their lost profits from infringers. By tightening up judge-made rules relating to proximate cause and the like, judges could achieve precisely this end. While this step may not go very far toward improving the system, it would encourage some acts of infringement that, under this hypothesis, would be socially beneficial. As a general matter, however, we think that this sort of approach would be a bad way of attempting to correct a flawed system. For one thing, the perceived merits or demerits of the present system are likely to vary substantially from one observer to another. Any suggestion that courts should manipulate the rules of enforcement and procedure to attain desired substantive goals ultimately will be futile, if courts themselves are divided on the issue of whether IPRs are already too strong or too weak. More fundamentally, however, this view threatens to undermine the rule of law by substituting judges' idiosyncratic views of the merits of (say) the patent system for the view expressed by Congress in enacting the Patent Act. To be sure, we are not so naive as to believe that the only thing judges do is to follow the rules laid down or that Congress in enacting the Patent Act and its many amendments has been motivated exclusively by concern for the public interest. Judges are policymakers and Congress often responds more to interest-group pressure than to considerations of the public good (to say nothing of the difficulty of ascribing a motive to a collective body, such as Congress, at all). But there is an institutional concern in operation here. Warts and all, legislative bodies are probably the best place for the fundamental decisions about the scope and duration of IPRs to be debated and resolved (barring some constitutional constraint upon legislative power, as may occur when IPRs come up against principles such as freedom of speech). We believe that, as a general matter, courts *ought* to operate as if the intellectual property laws embody the proper balance when deciding how to best calculate damages or craft standing and joinder rules, or the like, and leave it to other branches of government to decide whether the underlying assumption is true or false. That is the premise, at least, upon which this book is based.

The Law and Economics of IPRs

In this chapter, we provide an overview of basic patent, trade secret, copyright, and trademark law, and a general sense of the ways in which courts enforce these rights. Our principal focus will be on U.S. law, although from time to time we will examine other countries' laws and how they sometimes differ from U.S. law. With respect to each of these four bodies of law, we first provide a brief description of the legal rights at issue, and then follow with a discussion of the standard economic justifications for, and challenges to, these rights. Finally, we review the debate over whether intellectual property rights (IPRs) are better protected by property or liability rules.

PATENTS AND TRADE SECRETS

Inventions and other industrial know-how sometimes may be subject to ownership under patent or trade secret law.¹ Because the same invention may not be protected by both patent and trade secret law, and because patents usually confer a more robust form of protection, an inventor will usually choose patent over trade secret protection when either is available. In this section, we examine the scope of these bodies of law and their suggested economic underpinnings.

To be precise, inventions are protectable under the law of utility patents. Novel and distinct plant varieties are patentable in the United States for the same twenty-year period applicable to utility patents, under the Plant Patent Act, 35 U.S.C. §§ 161–4, and the Plant Variety Protection Act, 7 U.S.C. §§ 2321–2582. Plants also may be protected under utility patent law. See *J.E.M. AG Supply, Inc. v. Pioneer Hi-Bred Int'l, Inc.*, 534 U.S. 124 (2001). Designs that satisfy the three conditions of novelty, ornamentality, and nonobviousness are protectable for a fourteen-year period under the law of design patents, see 35 U.S.C. §§ 171–3. Other countries have analogous laws extending patent-like or *sui generis* protection to plant varieties and designs. Our discussion in this section centers upon the law of utility patents.

PATENT LAW

To qualify for patent protection, an invention must fall within the scope of patentable subject matter (a machine, process, manufacture, or composition of matter) and must meet the three statutory criteria of novelty, utility, and nonobviousness.² In the United States, the *novelty* requirement is normally satisfied as long as the patent applicant was the first to invent the claimed invention. For example, suppose that you file an application for a U.S. patent on a composition of matter comprising four elements: (A) water, (B) sugar, (C) electrolytes, and (D) glycerol in a concentration of approximately 0.5% to 5.0%. If, prior to the date on which you are deemed to have invented this composition, someone else already had invented and publicly disclosed in the United States a composition comprising these four elements, your invention would lack novelty.³ The *utility* condition requires only that the invention work and that it serve some minimal human need. Although utility is a minimal criterion, it does manage to weed out a few purported inventions, including those that cannot work (such as perpetual motion machines) and those that do not have a sufficiently specific known

² See 35 U.S.C. §§ 101, 103. Patent protection is not available for laws of nature, natural phenomena, and abstract ideas. See *Diamond v. Diehr*, 450 U.S. 175, 185 (1981). In practice, the scope of patentable subject matter has expanded in recent years, as U.S. courts have concluded that genetically engineered life forms, refined or isolated versions of naturally occurring physical substances, computer software-related inventions, and even business methods are all potentially patentable.

³ To be more precise, a claimed invention lacks novelty if another invention contains all of the “elements” or “limitations” of the claimed invention that are arranged in the same order and, prior to the date on which the applicant invented the claimed invention, the other invention was, *inter alia*, already known or used by others in the United States, or patented or described in a printed publication in the United States or abroad, or described in a pending and subsequently granted U.S. patent application, or was made and used in the United States by another who had not abandoned, suppressed, or concealed it, see 35 U.S.C. § 102(a), (e), (g). See also *id.* § 102(b), which denies patentability where the invention was patented or described in a printed publication, or in public use or on sale in the United States, more than one year prior to the date on which the applicant filed its U.S. patent application. Note that the invention lacks novelty (is anticipated) only if a single piece of qualifying prior art contains all of elements found in the invention. If you would have to put two or more pieces of prior art together to create an invention containing all of the elements, your invention might (or might not) fail the nonobviousness hurdle, but it would be novel.

Most other countries award the patent to the first to file a patent application, rather than the first to invent. In addition, many impose an absolute novelty rule under which an inventor is not entitled to a patent if the invention disclosed in the application was publicly disclosed anywhere in the world prior to the date of his application. The merits of a first-to-file or first-to-invent system, and of an absolute novelty rule, are beyond the scope of this chapter.

use; a possible example of the latter are intermediate research tools such as expressed sequence tags used in biotechnological research.⁴ The *nonobviousness* requirement denies patentability if the differences between the claimed invention and the relevant prior art are such that the claimed invention would have been "obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains."⁵ Nonobviousness is often the most difficult of the three conditions to satisfy. It is also the most difficult to describe or quantify, though in a rough sense it means that an invention is not patentable if it is an insubstantial improvement over the existing state of the art. In comparison with our previous example, a piece of prior art comprising (α) water, (β) sugar, (γ) electrolytes, and (δ) glycerol in a concentration of approximately 6.0% to 3.0% would not anticipate your invention – element D and element δ are not identical. Given the proximity of the range of glycerol concentrations, however, the prior art might render your invention obvious, unless (for example) the prior art "teaches away" (i.e., would lead the ordinary researcher to a different solution to the problem) from a lower glycerol concentration or the use of a lower concentration has unexpected properties.

Patent laws also impose upon the patent applicant a variety of disclosure requirements. Under U.S. law, the specification portion of the patent must include a written description of the invention "in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains . . . to make and use the same,"⁶ and also must disclose the inventor's own "best mode" or preferred embodiment of the invention as of the time the application is filed.⁷ It must "conclude with one or more claims particularly pointing out and distinctly claiming the subject matter that the applicant regards as his invention."⁸ Other countries' laws have similar, though not necessarily identical, requirements.

⁴ For more detailed discussions of utility, see *Brenner v. Manson*, 383 U.S. 519 (1966); 1 Chisum (2002; § 4.01, at 4–2.1); Schlicher (2002; § 3.02[1]); and see also Holman & Munzer (2000; 757–60). Most other countries have an analogous requirement, namely that inventions be "capable of industrial application."

⁵ 35 U.S.C. § 103(a). Most other countries state that the invention must demonstrate an "inventive step." The precise application of this third requirement varies somewhat from one patent system to another.

⁶ See 35 U.S.C. § 112. The portion of the statute just quoted actually contains two separate requirements: the written description requirement, which means that the description portion of the patent must conform to the claims and the enablement requirement, which means that the description must enable the hypothetical person of ordinary skill in the art to make and use the invention.

⁷ See 35 U.S.C. § 112.

⁸ *Id.*

Once the patent is granted, the U.S. patentee may exclude others from, among other things, making, using, or selling the invention in the United States for a term ending twenty years from the date on which the application was filed.⁹ If the patent owner suspects that someone is making, using, or selling without her permission, she can file suit for patent infringement. Infringement itself comes in two forms. First, the patent owner may claim that the defendant has *literally infringed*, by making, using, or selling an invention that contains all of the elements of the patented invention. In our previously mentioned hypothetical, where we used the letters A, B, C, and D to denote the elements of a patented invention, the defendant would literally infringe if he made, used, or sold a composition containing those same four elements (alone, or in combination with another element or elements). Alternatively, the patent owner may assert that the defendant's product (sometimes called the *accused device*) infringes under the *doctrine of equivalents*, which is substantially though not literally the equivalent of the patented invention. In our hypothetical, an accused device comprising elements A, B, C, and E would not literally infringe but might infringe by equivalents if the substitution of element E for element D is an insubstantial or trivial variation over the patented invention. Not surprisingly, applying the doctrine of equivalents can be quite complicated. Interpreted too broadly, the doctrine could have a chilling effect on follow-up inventors; interpreted too narrowly, it could render patents virtually worthless to the extent that almost any knowledgeable researcher could avoid literal infringement by making some minor modification to the patented invention.¹⁰

⁹ See 35 U.S.C. §§ 154(a), 271(a). For example, if you file your patent application on May 1, 1999, and the application is granted on May 1, 2001, the patent term begins on the latter date and ends twenty years from the former date (i.e., on May 1, 2019). Effective patent life may be shorter than the eighteen-year period in this hypothetical case, however. Patent owners often fail to pay modest maintenance fees, thus allowing their patents to lapse prematurely. And some products (such as pharmaceuticals) cannot be marketed until other government agencies, such as the Food and Drug Administration, approve them, and this process sometimes takes several years. In such cases, though, it is sometimes possible to obtain an extension of the patent term.

Note that patent ownership is defined in terms of the right to exclude others from making, using, or selling and not in terms of an affirmative right to make, use, or sell an invention. As noted in the preceding paragraph, other laws (such as food and drug laws) may preclude the patentee from practicing her own invention for a period of time. Or the patentee may have obtained a patent on an improvement to another's patented invention. For example, suppose that you have a patent on A, B, C, and D and that we obtain a patent on an improvement comprising A, B, C, D, and E. Neither we nor you would be entitled to practice the improvement without obtaining the other's permission. This phenomenon is referred to as the "blocking patents" problem.

¹⁰ For discussions of the U.S. law on the doctrine of equivalents, see *Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co.*, 535 U.S. 722 (2002); *Warner-Jenkinson Co. v. Hilton Davis*

The defendant in a patent infringement action usually offers two defenses: roughly, “I didn’t do it” and “It doesn’t matter.” The defendant asserts that he didn’t do it by denying that his conduct infringes, either literally or by equivalents. Second, the defendant is likely to argue that, even if his product comes within the scope of the patent claims, it doesn’t matter because the patent is invalid. A patent might be invalid for failing to satisfy any of the criteria of patentability, including novelty, utility, nonobviousness, and compliance with the relevant disclosure obligations; for example, the Patent Office might have overlooked, or misinterpreted the relevance of, some prior art that renders the invention obvious. The invalidity argument is successful in a substantial plurality of the litigated U.S. cases,¹¹ which suggests that the U.S. Patent and Trademark Office grants a large number of invalid patents. We return to this point in Chapter 7. A third possible defense is the *exhaustion* or *first-sale* doctrine. Once the patent owner releases into the stream of commerce a product that incorporates her patent, she cannot prevent the owner of that lawfully made product from using or reselling it, although she can prevent him from re-making it. As one would imagine, there is a fine line between lawful repair (a type of use) and unlawful remaking.¹²

If none of these defenses is successful, the defendant usually loses, because there are very few other exceptions to liability for patent infringement.¹³ Indeed, people usually describe patent infringement as a strict liability offense, and this description is more or less correct, though subject to a few caveats we raise in Chapter 5. To illustrate, suppose again that you own a patent on an invention comprising elements A, B, C, and D, and you discover that

¹¹ *Chem. Co.*, 520 U.S. 17 (1997); Conigliaro et al. (2001); and Wagner (2002). For discussions of the doctrine of equivalents in other countries, see, e.g., Weston (1998) and Yamamoto & Tessensohn (1999).

¹² Allison & Lemley (1998; 205–7) found 46% of the patents were invalidated in a population that consisted of all final written decisions on validity from 1989 to 1996; Moore (2000; 391) found invalidation rates of 29% and 36% in a sample that consisted of all the U.S. trial court decisions from 1983 to 1999 with, respectively, juries and judges deciding the issue of validity. Invalidity is raised as a defense in most patent cases.

¹³ For an overview of U.S. law, see Chisum (2002; § 16.03[3]). There is, for example, nothing as expansive as the copyright doctrine of fair use, which we discuss subsequently. In the United States, there is a very limited experimental use defense. Another exception exempts some conduct undertaken in connection with the submission of information to the Food & Drug Administration. See 35 U.S.C. § 271(e). Some countries recognize other limited exceptions, such as exceptions for private noncommercial use. This defense is not recognized in the U.S., as we discuss in Chapter 6. And in some rare instances a court may excuse literal infringement under the so-called *reverse doctrine of equivalents*. Under this doctrine, a court will permit the defendant to market an invention that literally infringes, but that is a radical improvement over the patented device. See Dam (1994; 266–7) and Lemley (1997; 1042–72).

we are making, using, or selling an invention comprising those same four elements without your permission. You sue us for patent infringement, but we assert that we did not copy from you; rather, we independently invented the same invention that is covered by your patent. Our assertion may be factually true – independent discovery does happen – but it is no defense to a charge of patent infringement. The laws of the United States and most other countries reduce the incidence of independent discovery to some extent by publishing most pending patent applications eighteen months after the date of filing; they also publish all issued patents, once the decision to grant the patent has been made. Patents therefore are public records, and so in theory the potential infringer could have discovered the patent in time to avoid the infringement. Again, we discuss the relevance of this observation in Chapter 5.

In a suit for patent infringement, the U.S. Patent Act (§§ 283–4) authorizes the court to award the prevailing plaintiff injunctive relief, as well as

damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer, together with interest and costs as fixed by the court.

“Damages adequate to compensate for the infringement” may include an award of the plaintiff’s lost profits attributable to the infringement; the amount of an established royalty; or a reasonable royalty. Significantly, U.S. courts do not interpret the Patent Act to provide for an award of the defendant’s profits attributable to the infringement, although patent law once permitted these restitutionary awards.¹⁴ A few other countries, including Canada and the United Kingdom, do permit restitutionary recoveries in patent law, but such recoveries appear to be rare.¹⁵ In addition, U.S. courts may award multiple damages up to three times the plaintiff’s actual damages, but courts generally exercise this discretion only in cases of willful infringement or bad faith litigation.¹⁶ The statute also permits the court to award attorney’s fees to the prevailing party in exceptional cases, which usually means those in which either the defendant is found to have willfully

¹⁴ As far as we can discern, when the U.S. Congress changed the law in 1947 to foreclose restitutionary awards, it was acting upon the perception that the cost of calculating these awards (in terms of complexity, cost, and delay) outweighed their benefits and that sometimes smaller competitors simply gave in rather than incur the cost of litigating. Given the high cost of patent litigation even without these awards, and given the need to perform a similar calculation in order to determine the amount of a reasonable royalty, we doubt whether this rationale makes much sense.

¹⁵ See generally Coury (2003).

¹⁶ See 35 U.S.C. § 284; 7 Chisum (2002; § 20.03[4], at 20–300).

infringed or the plaintiff obtained the patent by fraud or brought the action in bad faith.¹⁷ Prejudgment interest is usually awarded to the prevailing plaintiff as a matter of course.¹⁸

PATENT POLICY

The fundamental premise of the patent system is that society benefits when people conceive of new inventions; develop and commercialize new products incorporating those inventions (a process referred to as innovation, as distinct from invention); and publicly disclose information about their inventions, so that others may learn from and improve upon those inventions. Most people probably agree with this premise, and we will not bother to defend it. The difficult question is how to maximize these social benefits – or, more precisely, the surplus of social benefits over social costs. The patent system can be thought of as one way of attempting to achieve this goal.

To understand how the patent system may work to attain this end, it is important to recognize that one of the great things about information is that it is (or at least tends to be), in the language of economics, both *nonrival* and *nonexcludable*. Most tangible things are rivalrous, meaning that only one person or a small number of people can use or consume a particular good at any one time; while you use your computer or drive your car, for example, no one else can use that same computer or drive that same car. By contrast, a silicon chip located within your computer or your car may embody an invention that is simultaneously being used by thousands or even millions of people, and yet despite their simultaneous use the invention itself is never depleted (worn out). To put it another way, a nonrivalrous invention may be embodied in a rivalrous physical good. Only one person may be able to use that precise physical good at any one time and his or her use eventually may deplete that physical good. But the invention itself is an intangible thing, which is not depleted by use. The invention can be embodied in any number of physical goods.

Similarly, most goods are excludable, in the sense that you can take precautions – locks, guards, fences – to prevent other people from having access to them. But the only way to exclude others from having access to your idea for a new invention is to keep the idea to yourself. Once you disclose the idea to someone else, there may be no way to prevent that person

¹⁷ See 35 U.S.C. § 285; 7 Chisum (2002; § 20.03[4], at 20–384 to –385).

¹⁸ See 7 Chisum (2002; § 20.03[4][a], at 20–274 to –275). In the United States, there is no criminal penalty for patent infringement. There is in some countries.

from using it, assuming that he has the technical skill to do so. Of course, you could disclose your idea only to people you trust – or who are willing to sign nondisclosure agreements – and use locks and fences to keep everyone else from learning about it. But these measures are not foolproof. Once you embody your idea in a tangible object (say, a computer chip) and make copies of that object available to others (for example, by selling to them), the cat may be out of the bag. If the invention is valuable enough, someone will try to reverse-engineer it.¹⁹ In this way, knowledge of useful information tends to spread. In general, this dissemination of knowledge is desirable, because it enables others to use the invention and to improve upon it.

We said previously that one of the great things about inventions is that they are nonrivalrous and nonexcludable, but nonrivalrousness and non-excludability present a double-edged sword. Precisely because other people may be able to use your invention without your consent; once you have publicly disclosed it, your *ex ante* incentive to invest in creating or publicizing the invention is lower than it otherwise would be. Inventing something new often requires a substantial investment of time, money, and other resources. In the case of new drugs, for example, it may take hundreds of millions of dollars to come up with a safe, effective, and marketable drug; along the way, many promising candidates will be weeded out because of side effects or other problems. To copy someone else's new invention often costs considerably less, even if we factor in the cost of reverse-engineering. Thus, there may be a substantial incentive to take a free (or at least less costly) ride on someone else's investment. This potential for free-riding reduces the incentive to invent something new, because the inventor may be unable to recoup her sunk costs of invention. Competition from free-riders may reduce prices such that the cost of discovery and commercialization cannot be recovered. Moreover, to make optimal use of an invention, the inventor may need to disclose it to someone else who is better positioned to manufacture or market a tangible product that embodies the invention. But once she does so, the other party need not compensate her for the information; ideas, as we have said, tend to be nonexcludable. Of course, the parties could try to contract around this problem, but the potential recipient of the information may be unwilling to commit to not using the information until he knows what it is. After all, the information could turn out to be something in

¹⁹ Some things are easier to reverse-engineer than others. The formula for Coca-Cola has proven notoriously difficult to reverse-engineer. Decompiling computer source code from the underlying object code is difficult, though it can be done by skilled computer engineers. Other products, however, may be relatively easy, even obvious, for a person of average skill to unravel.

rse,
ling
ery-
nce
take
the
will
ble,

that
on-
ople
have
zing
new
ces.
ons
ong
side
osts
hus,
ride
the
e to
lucce
t be
may
ture
does
s, as
y to
tion
ows
g in

a has
from
teers.
skill

the public domain. But once the recipient knows what the information is, the disclosure already will have taken place. Nobel laureate Kenneth Arrow noted this “information paradox” in a famous essay some forty years ago.²⁰

The preceding analysis suggests that inventions are what economists refer to as a public good, similar to such things as parks, roads, national defense, and education. Almost everyone benefits from these sorts of things, either directly or indirectly. Even so, it can be difficult to limit access to public goods only to those who are willing to pay their fair share, whatever that means. Every beneficiary would be financially better off if he were to take a free ride, that is, to enjoy the benefits while others paid. If everyone acted this way, however, the good would not exist and everyone would lose out. In the case of parks, roads, national defense, and education, governments often try to solve this collective action problem by providing the public good and imposing taxes (a sort of user fee) upon the beneficiaries (in theory, everyone), to pay for it.²¹

For similar reasons, the free-rider problem may undermine the incentive to create, disclose, and commercialize new inventions, absent some corrective measure. Needless to say, the possibility of free-riding will not *always* have this effect; people were coming up with new inventions long before patent law ever came into existence, for a variety of reasons. Some inventions may not entail substantial sunk costs. Some inventors may be able to make up those costs merely by being the first on the market with a new product (the so-called first-mover advantage), particularly if the product is difficult to reverse-engineer or the cost of copying otherwise remains high. (It is no coincidence that intellectual property laws made little headway until the cost of copying started to come down, in the seventeenth and eighteenth centuries, due to the printing press and other innovations.) In other cases,

²⁰ Arrow (1962; 615).

²¹ Attempting to solve the public goods problem through taxation gives rise to its own set of familiar problems. Some people will try to avoid paying their fair share; there will be debates over what a fair share is; interest groups will lobby government for special favors or exemptions; and so on. Ideally, a policymaker would consider the costs and benefits of other options for solving the free rider problem, as Coase and others remind us. See Coase (1974). Indeed, the law of intellectual property generally does not follow the taxation solution, for reasons we discuss in the text previously, although government does provide direct financing and tax benefits for some scientific research. In addition, some commentators have recently argued that a system under which government sometimes conferred prizes, or effected buyouts of patents, might be superior to the patent system alone as a way of addressing the public goods problem. See Abramowicz (2003), who reviewed the literature and proposed his own solution. For now this remains an interesting theoretical alternative, but it is not directly relevant to our exploration of optimal remedies under the current system.

self-help measures such as locks, fences, or nondisclosure agreements may be sufficient to deter copying until the sunk costs have been recouped. In yet other cases, private investors, philanthropists, or potential beneficiaries could promise *ex ante* to pay inventors for developing new inventions or could offer the inducement of *ex post* rewards and prizes. Nevertheless, the intuition remains that, absent further measures, some inventions that would benefit society would not be invented or would not be optimally disclosed or commercialized. The patent system is premised on the reasonable assumption that the public will enjoy additional benefits when the government takes additional steps to encourage the creation, commercialization, and disclosure of new inventions. More things will be created, commercialized, and disclosed. Also, some of the costs of self-help measures, such as locks and fences, can be avoided.

As we noted earlier, the way the law handles many other public good problems is through taxation: that is, the government provides the good and pays for it from tax revenues. Sometimes governments employ this method or analogous methods to finance the production of new inventions as well. Governments use tax revenues to fund some scientific research, and they also grant some tax benefits to private firms that engage in R & D. One problem, however, with relying exclusively or predominantly on these methods is that governments, unlike the decentralized marketplace, may lack the knowledge of what needs to be invented or how to value new inventions. And unlike most private actors, governments may not have to submit to the discipline of the marketplace. Incorrect decisions on the part of government decisionmakers could result in insufficient investment in invention, or investing in the wrong types of invention, or under- or overvaluation of that which has been invented. Competition for government funding also could lead to familiar rent-seeking²² problems on the part of private actors or agency capture by the affected industries.

All of these considerations lead most theorists to conclude that a different method for encouraging invention – namely, the one a patent system provides – is preferable to relying predominantly upon direct funding. Under a patent system, society “funds” invention by allowing private actors to

²² “Rent-seeking” behavior occurs when people seek economic “rent,” that is, value in excess of that which they could obtain from their next-best investment. (More technically, economic rent is the value derived from an investment in excess of one’s opportunity cost.) Some rent-seeking behavior may be socially inefficient (e.g., competition to obtain privately valuable, but socially wasteful, government benefits). In other instances, competition to obtain economic rents – such as the potential profits to be derived from a valuable patent – may dissipate the value of those rents. See *infra* pages 18, 20.

decide what to invent and then conferring exclusive rights upon them for a limited time. These exclusive rights are worthless if the invention turns out to be a dud, but ultimately the market decides what is valuable and what is not. If the inventor (or her licensee) can come up with a marketable embodiment of the invention, the exclusive rights will provide the inventor with an opportunity to recover her sunk costs. Correcting for the free-rider problem in this manner is the genius of the patent system.²³

Critics nevertheless note that patents come with some substantial social costs and sometimes charge that the benefits may be overstated as well. As for the benefits, the empirical evidence that patents provide a necessary incentive to inventive activity is hardly overwhelming. Survey evidence suggests that many firms rely more heavily upon other incentives to invent (such as first-mover advantages and trade secrecy) and also that firms often patent for strategic reasons, such as preventing others from gaining a competitive advantage.²⁴ Patents nevertheless may provide a substantial incentive to invent in certain industries, such as pharmaceuticals, where sunk costs are particularly high. Their assumed role in inducing disclosure and commercialization might be important even if the incentive to invent is of lesser magnitude.

Patent systems do give rise to a variety of social costs, however, which policymakers must consider and which an ideal patent system would try to minimize. One is the systemic cost of processing, enforcing, and maintaining patent rights, which requires at a minimum a patent office and courts to resolve patent infringement claims. This cost is not terribly significant in the industrialized nations, but may prove burdensome for some less developed countries. Second is the potential for the patent system to inhibit future invention or innovation that is based upon existing patented inventions. Invention tends to be cumulative and thus one consequence of a patent system is to raise the cost of creating follow-up inventions based upon an earlier technology.²⁵ This could prove problematic, even if the follow-up inventor is willing and able to pay for permission to use the patent (which may not always be the case, due to budget constraints, the uncertainty of the future payoff, and so on). If the follow-up inventor must negotiate

²³ For further discussion of the traditional arguments that patents induce invention and disclosure, see, e.g., Dam (1994; 247) and Scotchmer (1991; 31); see also *Aronson v. Quick Point Pencil Co.*, 440 U.S. 257, 262 (1979). For a focus on the incentive to commercialize, see Kieff (2000) and Turner (1998; 186–93).

²⁴ See Cohen et al. (2000; 9–11 & Figures 1–4); Levin et al. (1987; 794–5); Mansfield et al.

(1981; 915); Scherer et al. (1959; 118); and Taylor & Silbertson (1973; Chapter 9).

²⁵ See Merges & Nelson (1990) and Scotchmer (1991).

with a multiplicity of previous researchers, the mere cost of transacting could be enormous; critics argue that this problem is becoming acute in fields such as biotechnology,²⁶ although this critique is hotly disputed.²⁷ A third social cost is the cost associated with duplicative effort: the potential availability of a patent may induce a “race” to become the first to invent. At worst, such attempts to obtain patent “rents” can result in substantial wasted effort, because only one firm can win the race, and the expected benefits of obtaining the patent may be dissipated in the effort to win the race. On the other hand, patent races may speed up the development of new technology and may have unintended, but beneficial, spillover effects: on the way to discovering one thing, a researcher may discover something else that was unforeseen.²⁸ In light of these conflicting effects, the theoretical literature on patent races remains equivocal.

A fourth cost is the potential cost associated with monopoly rights. Monopolies can be troubling for two different reasons. First, monopolies transfer wealth from consumers to monopolists and depending on the circumstances and on one’s theory of distributive justice, this outcome may be undesirable. Second, a monopoly is, in economic terms, allocatively inefficient, because it reduces social welfare. In other words, it is not just that the monopolist gets one more slice of the pie and consumers one less slice; the size of the pie is smaller. A monopolist maximizes profit by producing where marginal cost equals marginal revenue – we depict this phenomenon graphically in Chapter 3 – with the result that (compared to the outcome under perfect competition) price increases while both output and social welfare decrease. The monopolist is better off, but consumers are worse off and consumers’ losses outweigh the monopolist’s gains. Economists refer to this decrease in social wealth as a “deadweight social welfare loss.”

Although the deadweight loss is a potential cost of any system of exclusive rights, it is important to bear in mind that most patents do *not* confer monopoly rights in any economically meaningful sense (despite the fact

²⁶ See Heller & Eisenberg (1998).

²⁷ See Wagner (2003; 12–13) for a review of empirical and theoretical challenges to the Heller-Eisenberg thesis.

²⁸ For discussion of some of the relevant literature, see Abramowicz (2003; 183–8); see also Scotchmer (1998; 275), who noted “two views on patent races: that they inefficiently duplicate costs and that they efficiently encourage higher aggregate investment”; see also Tirole (1988; 400), who noted that the loser in a patent race may benefit from positive spillovers, may develop another product, and may gain experience for future races. Merges & Nelson (1990; 870–9) argued that empirical evidence is more consistent with the theory that competition in the market for improvements spurs innovation, despite possible efficiency losses attributable to rivalrous invention. See also Reinganum (1989; 853–68).

icting
ite in
l.²⁷ A
ential
at. At
asted
its of
n the
ology
ay to
t was
ature

Mo-
rans-
cum-
ay be
inef-
that
slice;
icing
enon
some
ocial
e off
er to
exclu-
onfer
fact

eller-
e also
duplic-
Tirole
overs,
elson
com-
losses

that people often carelessly refer to patents, and sometimes other IPRs as well, as monopoly rights). For one thing, most patented inventions “read on” (that is, cover or render infringing) components of larger products and not on discrete products themselves. Often the market for the discrete product will be competitive, even if only one producer has the right to manufacture a particular component. A related point is that for most patented inventions (whether components or discrete products) there is a range of acceptable, nonpatented substitutes, which limit the patent owner’s ability to obtain a monopoly profit. Indeed, a majority of patents are never commercialized at all, much less meet with commercial success. Of course, a few patents *do* result in the lucky patent owner obtaining a temporary monopoly or at least some competitive advantage even if it falls short of full-blown monopoly power. Indeed, the whole point of patent rights as we have outlined them is the potential they hold for the inventor to price above marginal cost for at least as long as is necessary to recoup the sunk costs of invention. Certainly patents do fulfill this promise sometimes. When they do, however, one must reckon with the corresponding deadweight loss. Even in this instance, however, one must consider how the world would have looked in the absence of the patent system. Even if the invention leads to temporary monopoly power, consumers might be better off than if the invention had not been invented at all or had been invented much later in time.²⁹

There are some additional theories of patent rights that complement the standard incentive theory.³⁰ Of these, the most well known is Edmund Kitch’s “prospect” theory. As developed by Kitch, this theory holds that patent rights enable inventors to efficiently coordinate investments by others in second-generation improvements.³¹ Kitch argues that “pioneering” inventions in particular – meaning those that are likely to have a large number of follow-up applications – merit a broader patent “scope” than do more pedestrian inventions, because the pioneer patent owner can reduce the amount of rent-seeking by potential improvers. Kitch’s theory remains influential but has been critiqued on several grounds. For one thing, the strong patent rights that the prospect theory appears to contemplate may weaken the incentive to create follow-up improvements, because the improver will

²⁹ See Arrow (1962; 619–20); Dam (1994; 251).

³⁰ In addition to the literature on the prospect theory discussed subsequently, the interested reader is advised to consult Kieff (2003), who argued that a number of patent doctrines function to reduce administrative costs; Long (2002) developed a theory that patents signal positive firm attributes.

³¹ See Kitch (1977; 267–71).

capture only a portion of the gains from improvement. (On the other hand, allowing the follow-up improver to patent his invention provides him with some leverage, as does the possibility that the improvement will be held noninfringing under the reverse doctrine of equivalents.³²) The prospect of obtaining broad patent rights also may stimulate races to obtain the pioneering patent with the potential negative consequences described earlier.³³ The prospect theory remains empirically suspect as well. One historical study argues that the development of follow-up improvements in fact has often had little connection with the existence or nonexistence of patent rights in the original discovery; another argues that patents owners often have not used broad patent rights to coordinate follow-up innovation, but rather have engaged in satisficing behavior.³⁴ Nevertheless, the prospect function remains a prominent theoretical justification for the patent system, even if only in conjunction with the more mainstream incentive theory.

In summary, the previous arguments posit that the ideal patent system would maximize the surplus of social benefits over social costs. The benefits include the development of new inventions, as well as their disclosure and commercialization (and, possibly, the prospect function described earlier); the costs include the various administrative, transaction, monopoly, and rent-seeking costs we have discussed. How close any patent system comes to satisfying this ideal, in comparison with alternative methods, remains a matter of speculation and will depend in part upon how "strong" or "weak" the patent rights under consideration are. Patent strength is a function of both duration and scope. Stronger rights – rights that have longer duration or broader scope – may increase the incentive to create, disclose, innovate, and coordinate investment in follow-up improvements, but they also may increase the attendant social costs.

The effect of patent duration upon patent strength is easy to understand: the longer the patent lasts, the more potential value it has and vice versa

³² See Footnotes 9 and 13 in this chapter.

³³ See McFetridge & Smith (1980). Grady and Alexander argue, however, that the patent system limits rent-seeking in various ways. For example, it does so by foreclosing patent protection for pioneering discoveries that cannot be improved upon, such as laws of nature. See Grady & Alexander (1992). For a critique of the Grady & Alexander thesis, see Merges (1992).

³⁴ See Beck (1983) and Merges & Nelson (1990; 871–8). The term "satisficing" was coined by Nobel laureate Herbert Simon. Simon posited that firms sometimes seek not to maximize profits, but rather to attain a certain level of satiation: "a certain level or rate of profit... a certain share of the market or a certain level of sales. Firms would try to 'satisfice' rather than to maximize." Simon (1959; 263).

ind, with held it of pio- er.³³ ical has tent often but ect tent tive

tem efits and er); and nes is a ak" 1 of ion ate, nay

nd: ersa

tent tent ure. rges

1 by nize ... a ther

(although in practice most patents are obsolete long before their expiration dates).³⁵ As we noted, the patent term for all inventions is more or less standardized and runs from the date of grant until twenty years after the date of filing. In theory, there is no reason why this must be so. One could, for example, award longer patents for inventions that are the result of large investments in research and development, or that confer substantial social benefits, and shorter patents for other inventions. Modern patent systems nevertheless follow this approach in favor of a "one-size-fits-all" patent term.³⁶ Arguably, this approach is justified in light of the difficulty of properly (and neutrally) determining the appropriate duration of patent rights; in any event, the worldwide standard today is a patent term ending twenty years from the date of the patent application with few exceptions.³⁷ Some inventions, therefore, may receive much more of a stimulus than is necessary, while others not enough.

Patent scope can refer both to patentable subject matter (and related issues) and to patent breadth. Patent scope in the sense of subject matter can be either broad or narrow. As discussed earlier, in the United States patent scope in this sense is quite broad, although even here it does not extend to the discovery of abstract ideas, laws of nature, and naturally occurring physical phenomena.³⁸ More generally, one might use the term patent scope to

³⁵ See Lemley (2001:1503–4) who showed that more than 2/3 of U.S. patents lapse prematurely due to the owners' failure to pay modest maintenance fees. In the United States, maintenance fees are due at the end of 3½ years, 7½ years, and 11½ years from the date of the grant in amounts that range from \$910 to \$3,220.

Some countries, however, also award protection for minor innovations under a system of protection known as "utility models," "petty patents," or other names. (Australia, for example, recently revised and renamed its petty patent system to a system of "innovation patents.") In general, countries with utility model protection make this protection available to inventions that otherwise might not qualify for patent protection (because they are insufficiently nonobvious, for example). These petty patents typically confer exclusive rights for a short period of time, such as five or ten years. The United States does not award utility model protection.

The international treaty known as the TRIPs Agreement, to which all member nations of the World Trade Organization (WTO) are parties, requires a minimum patent term ending twenty years from the date of application. It also generally forbids member nations from discriminating with respect to fields of technology, such as from awarding different patent terms for different types of inventions. Thus, whatever the theoretical merits may be of a more perfectly calibrated system, such a system does not appear to be a realistic possibility anytime soon.

³⁶ A system that awarded patents to the first person to discover a new law of nature would marginally increase the incentive to be such a discoverer. But the additional incentive may be unnecessary, in light of existing incentives such as research grants, prizes, and fame, and in light of the high social cost of conferring exclusive rights in such cases. See also Grady & Alexander (1992), who argued that conferring patent protection

delineate the class of inventions that fall within the definition of patentable subject matter *and* that satisfy the other requirements of patentability. For example, every patent system must decide what the terms novelty, utility, and nonobviousness mean within that system. In addition, every system will exclude from patent protection inventions that do not satisfy these requirements.³⁹ More commonly, however, when people use the term patent scope they mean patent breadth; a patent is broad if it reads on many possible embodiments. Patent breadth, in turn, is a function of several patent doctrines. Patent attorneys can draft individual patent claims broadly or narrowly. Drafting them too broadly, however, risks invalidation on a number of grounds, including enablement or anticipation by the prior art. In this sense, substantive patent law, by defining what an enabling disclosure is or what “counts” as prior art, indirectly controls patent scope. The law of infringement is yet another mechanism by which patent scope can be regulated. As we have seen, a patent confers the right to exclude others not only from literally infringing, but also from infringing by equivalents. A broad interpretation of the term “equivalent” expands patent breadth, whereas a narrow interpretation reduces it.

Patent scope is not quite as “one-size-fits-all” as is patent duration. In applying the doctrine of equivalents, for example, U.S. courts are directed to confer broader scope upon pioneering inventions.⁴⁰ Dan Burk and Mark Lemley also have argued that U.S. courts appear to have certain preconceptions – some of which may be incorrect – about various fields of technology, such as what sorts of inventions within those fields are obvious applications of the prior art, what type of information must be disclosed to enable a person of ordinary skill to practice the invention, and so on.⁴¹ To the extent this observation is true, it suggests that patent scope, in a broad sense, will vary to some extent from one field to another. This type of variation, however, arises from the courts’ and patent offices’ application of general rules designed for all patentable subject matter and not from industry-specific statutory provisions.

upon discoveries which cannot be improved upon, such as laws of nature, would result in patent-race rent dissipation with no offsetting benefits in terms of follow-up invention.

³⁹ Kieff argues that many of the relevant rules can be explained as an attempt to minimize the administrative costs of the patent system. See Kieff (2003).

⁴⁰ See, e.g., *Augustine Med., Inc. v. Gaymar Indus.*, 181 F.3d 1291, 1301 (Fed. Cir. 1999) where this dictum was noted, although it was tempered with the observation that pioneering inventions typically have broader scope even without benefit of the doctrine of equivalents because claims to such inventions are less constrained by the prior art.

⁴¹ See Burk & Lemley (2002).

table
For
ility,
stem
e re-
itent
pos-
tent
y or
um-
t. In
sure
w of
equ-
only
road
as a

ion.
cted
and
tain
elds
bvi-
osed
n.⁴¹
in a
type
rom

d re-
v-up
e the

here
ering
lents

scholars; therefore, recognize that patent rights can be strengthened (or diminished) by expanding (or reducing) scope, duration, or both. Two points nevertheless remain problematic. First, there is little theoretical and no empirical evidence addressing the issue of what the optimal scope or duration of patent rights should be in the real world. Indeed, it is not clear how such empirical evidence ever could be produced; controlled experiments do not appear feasible. A second problem is that no one knows much about the tradeoff between scope and duration (e.g., how much scope could be reduced if the duration were extended without affecting patent strength). Some theoretical literature *does* address this issue, however, as well as the related issues of whether it is less socially costly to confer patents of broader scope and short duration or of narrow scope and long duration.⁴² But no consensus has yet emerged as to which, if any, departures from the present system would increase social welfare.

TRADE SECRETS

Trade secret law differs from the law of patents in several crucial respects. One is that trade secret law is less uniform than patent law. In the United States, trade secret protection is based primarily on common law and state statutory law. Thus, unlike patent law that is based on a federal statute, trade secret protection can vary to some extent from one state to another. These differences have been reduced in recent years, however, by most states' passage of the Uniform Trade Secret Act (UTSA). Other countries' trade secret laws, nevertheless, may vary widely and there is very little international law addressing the topic.⁴³ When we discuss trade secrets in this book, our principal focus will be upon U.S. law as embodied in the UTSA.

A second difference is that trade secret protection is much easier to obtain than is patent protection. Under the UTSA, any information that provides a person with a competitive advantage as long as it remains secret is potentially protectable as a trade secret. The stringent novelty and nonobviousness conditions of patent law do not apply. Thus, even such unpatentable items as customer and supplier lists, recipes, and the amount of a secret bid can qualify as trade secrets; although potentially patentable, unpatented inventions

See Gallini (1992) who advocated shorter, but broader patents, and took issue with other theorists who have advocated longer and narrower patents; see also Ayres & Klemperer (1999; 987 Footnote 2).

TRIPs and NAFTA each contain one article dealing with trade secrets, but neither provides anywhere near as much specificity as the UTSA. Countries, therefore, retain considerable leeway to craft their trade secret laws as they see fit.

can qualify as trade secrets, too. Moreover, whereas patent law requires the patentee to disclose certain information to the public as a precondition of obtaining a patent, trade secret law affirmatively discourages the owner from making any public disclosure, because any such disclosure of trade secret information may result in the information losing its protectable status. Indeed, there is no government body, analogous to the patent office, that affirmatively bestows trade secret protection. Any information that qualifies as a trade secret under the definition above *is* a trade secret; no registration or other formalities are required.

From the standpoint of the inventor, the upside of trade secret protection is that it is much easier to obtain than patent protection; the downside is that it is also less robust and often more vulnerable to forfeiture. The owner of a trade secret may exclude another from, among other things, acquiring the secret by "improper means" such as theft or espionage or from using or disclosing the secret if the other knew (or had reason to know) at the time of disclosure or use that the secret was derived from a person who (1) had used improper means to acquire it, or (2) had acquired it under circumstances giving rise to a duty to maintain secrecy, or (3) owed a duty of secrecy to another.⁴⁴ To illustrate, suppose that your secret cookie recipe qualifies as a trade secret. Under the UTSA, you would have legal recourse against someone who breaks into your headquarters and steals the recipe. (Of course, you would also have recourse under other civil and criminal laws; occasionally, though, trade secret law penalizes an acquisition that would not violate any other body of law.⁴⁵) You also would have a claim against someone (for example, a current or former employee) who discloses the recipe to a recipient in violation of a duty of secrecy imposed by contract or by the common law of agency. (Again, you might have other causes of action as well for breach of contract or breach of fiduciary duty.) Finally, you might have recourse against the recipient, if for example the recipient uses the recipe despite the fact that she knows, or should know, that the disclosing party violated a duty of secrecy. Often, a trade secret claim will be the only recourse against the recipient who in our hypothetical does not appear to be in privity of contract with the owner and may not owe him any common law fiduciary duty.

Unlike a patentee, the trade secret owner has *no* recourse against independent discovery or reverse engineering. Moreover, trade secret protection

⁴⁴ See UTSA §§ 1(2)(i), (ii); Restatement (Third) of Unfair Competition § 40.

⁴⁵ See *E. I. du Pont de Nemours & Co. v. Christopher*, 431 F.2d 1012 (5th Cir. 1970).

s the
ition
wner
e se-
atus.
that
lifies
ition

tion
de is
vner
ring
sing
the
who
nder
duty
cipe
urse
sipe.
aws;
ould
inst
the
ract
s of
ally,
ient
the
will
not
any

ide-
ion

lasts only for as long as the information remains secret and valuable. Unfortunately, the information may become widely known despite the trade secret owner's best efforts to maintain secrecy and thus no longer be a secret; or it may become obsolete and, therefore, lose value. In these respects, trade secret protection is more tenuous and less valuable than patent protection, though one must take this observation with a grain of salt. On the positive side, trade secret protection may subsist in some subject matter – for example, customer lists and insufficiently nonobvious inventions – that would never qualify for a patent. In addition, because the duration of trade secret protection is indefinite, in some rare instances protection may persist for much longer than the term of a patent. Good examples include the formula for Coca-Cola and the recipe for Kentucky Fried Chicken. A related point is that obtaining a patent destroys the secrecy of the information, whereas trade secret protection does not. So, if the information is particularly difficult for others to reverse-engineer, trade secret protection can be more valuable than patent protection.

If the trade secret owner is able to prove the actual or threatened misappropriation of a trade secret, the court may award injunctive relief. The UTSA qualifies this right to an injunction, however, by providing that:

In exceptional circumstances, an injunction may condition future use upon payment of a reasonable royalty for no longer than the period of time for which the user could have been prohibited. Exceptional circumstances include, but are not limited to, a material and prejudicial change of position prior to acquiring knowledge or reason to know of misappropriation that renders a prohibitive injunction inequitable.⁴⁶

The plaintiff is also entitled to recover damages, which may include "both the actual loss caused by misappropriation and the unjust enrichment caused by misappropriation that is not taken into account in computing actual loss"; in the alternative, the court may award a reasonable royalty for a misappropriator's unauthorized disclosure or use of the secret.⁴⁷ Finally, in the event of a "willful and malicious" misappropriation, the UTSA permits

⁴⁶ UTSA § 2(b); see also Restatement (Third) of Unfair Competition, § 44 cmt. c. See UTSA §§ 3(a), (b); see also Restatement (Third) of Unfair Competition § 45 (similar). A damages recovery may be conditioned, however, on the defendant's not having incurred "a material and prejudicial change of position prior to acquiring knowledge or reason to know of misappropriation." UTSA § 3(a); cf. Restatement § 45 cmts. b and g, suggesting that the court may award a reasonable royalty for use made after the user is put on notice that the information is secret and an injunction conditioning further use upon payment of a royalty.

punitive damages in an amount not exceeding twice the amount of actual damages, as well as attorney's fees.⁴⁸

Commentators have suggested that trade secret law supplements the patent system by providing an incentive to develop information that has some social value, though not enough to warrant a patent.⁴⁹ Perhaps more importantly, trade secret law arguably discourages socially wasteful measures to protect the secrecy of one's invention. A trade secret owner is required to take reasonable precautions (such as building a fence) in order to maintain trade secret protection, but the law may make it unnecessary for him to build a one-hundred-foot-high fence.⁵⁰ Trade secret law departs from patent law, however, insofar as it discourages the public dissemination of information. To critics, this aspect of trade secret law is sufficiently problematic to call the entire body of law into question.⁵¹ The secrecy-enhancing character of trade secret law is nevertheless constrained to some degree by the rule permitting others to independently discover or reverse-engineer the secret, a point to which we return in Chapter 5.

COPYRIGHT

In 1710, the British Parliament enacted the first modern copyright law, the Statute of Anne, which thereafter became the model for the first U.S. Copyright Act and influenced early copyright legislation in other countries as well. As originally conceived, copyright in the United States and the United Kingdom subsisted only in "books, maps, and charts" and protected the author (or his assignee) only against unauthorized "printing, publishing, republishing, and vending."⁵² Protection lasted for fourteen years (the same as the original patent term) and was measured from the date of first publication, although it could be renewed for an additional fourteen years. In addition, courts granted protection only against the literal or near-literal copying of a work in substantially its entirety, thus permitting the publication of unauthorized abridgements, sequels, and even translations until well

⁴⁸ See UTSA §§ 3(b), 4; Restatement (Third) of Unfair Competition § 45 cmt. i, which permits "punitive damages under the rules generally applicable in the jurisdiction to the award of punitive damages in tort actions."

⁴⁹ See Friedman et al. (1990; 63–4).

⁵⁰ See Burk (1999; 173) and Note (1992).

⁵¹ Bone, for example, argues that the social benefits of trade secret law are modest, because much activity that trade secret condemns would violate other laws and that imposing trade secret liability upon other forms of conduct raises social costs without sufficient countervailing benefits. See Bone (1998).

⁵² See Act of May 31, 1790, Chapter 15, § 1, 1 Stat. 124, 124 (repealed 1802).

ctual
s the
t has
sures
ed to
ntain
build
t law,
tion.
ll the
trade
tting
nt to

law,
U.S.
oun-
l the
cted
lish-
(the
first
ears.
teral
lica-
well

mits
rd of

cause
using
cient

the nineteenth century. Over time, however, copyright protection has expanded to include many other works of authorship and additional rights have been added.

Today, copyright laws in the United States and elsewhere protect virtually all "original works of authorship," including literary,⁵³ musical, dramatic, and choreographic works; pictorial, graphic, and sculptural works; motion pictures and other audiovisual works; architectural works; and, in the United States and some other countries, sound recordings.⁵⁴ Originality is usually the only condition to satisfy; in the United States, it means only that the work exhibit independent creation and some minimal degree of creativity either in the expression of underlying facts or ideas or in the selection or arrangement of those facts.⁵⁵ Significantly, ideas and facts themselves are not subject to copyright protection. For example, the date on which the Red Baron died (a fact), as well as the various ideas scholars have proffered concerning who shot him down, must remain in the public domain no matter how much work went into discovering or formulating them. Not surprisingly, the line between protectable expression, selection, or arrangement, on the one hand, and unprotectable ideas or facts, on the other, is often difficult to discern; the issue often comes to a head in cases involving labor-intensive, but unoriginal, presentations of facts. When copyright exists, it subsists from the moment of creation⁵⁶ and vests in the author of the work.⁵⁷ The standard copyright

Literary works are works, other than audiovisual works, expressed in words, numbers, or other verbal or numerical symbols or indicia...." 17 U.S.C. § 101. They include both the source code and the object code of computer programs.

A musical work would include such things as songs, symphonies, and concertos. A sound recording is a work that results "from the fixation of a series of musical, spoken, or other sounds," such as might be embodied in a compact disk or cassette tape.

See *Feist Publications, Inc. v. Rural Tel. Serv.*, 499 U.S. 340, 344–51 (1991).

This means that copyright protection, like trade secret protection but unlike patent protection, exists without the need for any government agency to first pass judgment on whether the work qualifies for protection. This is a relatively new development in the United States. Prior to 1978, state common-law copyright existed from the moment of creation, but it terminated upon publication at which point federal copyright protection sprung into existence, but only if all published copies of the work bore the appropriate notice of copyright. Failure to comply with the notice requirement resulted in the forfeiture of copyright. Registration of copyright is still required as a precondition to filing a copyright infringement action in the United States, subject to certain exceptions. Usually, though, registration is a mere formality, unlike the process of obtaining a patent.

Oddly enough, there is no definition of the term "author" in the U.S. Copyright Act. In most countries, only the human being or beings who created the work can be authors, although exceptions are sometimes made in the case of motion pictures, computer software, and newspaper articles (in which instances employers sometimes are viewed as the owners of the work). In the United States, works created by an employee within the scope of his employment, as well as some specially commissioned works, are defined as "works made

term in the United States and most other industrialized nations now consists of the life of the author plus seventy years.⁵⁸

Although they now encompass much more than the rights of “printing, publishing, republishing, and vending,” the rights of a copyright owner remain somewhat less expansive than the corresponding patent owner’s rights. Of paramount importance is the reproduction right, that is, the right to reproduce protectable expression, selection, or arrangement in tangible copies.⁵⁹ This right protects not only against literal copying but, in appropriate cases, against copying such aspects of a work as its plot and its fictional characters – though deciding at precisely what point these aspects of the work fall on the “expression,” as opposed to the “idea,” side of the line can be quite difficult.⁶⁰ The copyright owner also has the exclusive right to prepare derivative works, that is, works that are based upon the copyrighted work (such as a translation or a motion picture version of a novel).⁶¹ This right, sometimes known as the adaptation right, is largely, though not entirely, coextensive with the modern reproduction right. In addition, the owner has the exclusive right to distribute copies of the work to the public; to perform and display the work publicly; and to import the work into the United States.⁶² Each of these terms is a term of art; although we will not

for hire,” copyright to which subsists *ab initio* in the employer or hiring party. See 17 U.S.C. §§ 101 that defines work made for hire, 201(b). Most countries, including the United States, also permit the author to freely assign or license most of the rights comprising copyright protection, subject to certain exceptions. In a few countries, however, authors can only license, but never assign, their copyrights.

⁵⁸ The U.S. Supreme Court recently upheld the life-plus-seventy term, which was adopted only in 1998, against constitutional challenge. See *Eldred v. Ashcroft*, 537 U.S. 186 (2003). The copyright term for works for hire and certain other works consists of ninety-five years from the date of publication or 120 years from the date of creation, whichever is shorter. Works published prior to January 1, 1978, but still under copyright protection in the United States as of January 1, 1999, enjoy a ninety-five-year term.

⁵⁹ See 17 U.S.C. § 106(1).

⁶⁰ See *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930) (Hand, J.) for details regarding the formulation of the famous “abstractions” test.

⁶¹ See 17 U.S.C. § 101 that contains the definition of derivative work, § 106(2).

⁶² See 17 U.S.C. §§ 106(4), 106(5), 106(6), 602. The performance right applies to literary, musical, dramatic, and choreographic works, pantomimes, motion pictures and other audiovisual works, and to digital sound recordings. See 17 U.S.C. §§ 106(4), 106(6). The display right applies to “literary, musical, dramatic, and choreographic works, pantomimes, and pictorial, graphic, or sculptural works, including the individual images of a motion picture or other audiovisual work.” *Id.* § 106(5).

Many countries also confer upon authors or other entities additional rights known as *moral* rights and *neighboring* rights. Moral rights laws typically entitle the author to claim authorship of his work and to prevent certain distortions and mutilations of the work, even after the title to the work and its copyright have passed to another. U.S. law incorporates

sists
ting,
vner
ier's
ight
ible
princ
onal
the
can
pre-
ited
This
en-
the
olic;
the
not

S.C.
ates,
ight
only

pted
03).
years
rter.
ited

tails

ary,
ther
The
nes,
tion

n as
aim
ven
ates

well upon it here, there is considerable case law addressing such knotty issues as whether a particular performance is “public,” whether a particular work qualifies as a “distribution,” and so on.

As in patent cases, defendants in copyright cases often argue that their product does not infringe or that the plaintiff’s IPR is invalid (for example, because the plaintiff’s work lacks originality). Courts have developed a variety of tests, all of them rather vague, for determining whether an accused work is “substantially similar” to the complainant’s work; we briefly refer to some of these in Chapter 5. In addition, copyright law admits many more exceptions than patent law. There is, first, a *first-sale* or *exhaustion* doctrine, similar to what we find in patent law, which permits the owner of a lawfully made copy to distribute and display that copy without permission of the copyright owner.⁶³ Second, a variety of limited exceptions apply only to certain works or certain uses; many of these provisions of the U.S. Copyright Act tend to be highly technical.⁶⁴ Third, the United States recognizes the *fair use* defense, an open-ended exception that, when successful, can exempt the defendant from liability in a number of different situations. Courts generally consider four factors in deciding whether a use is fair, and thus exempt; we briefly discuss this issue in Chapter 6. (Other countries generally eschew fair use in favor of more specific, less open-ended exceptions. Even the analogous “fair dealing” exception found in some countries of the British Commonwealth is considerably narrower than the fair use doctrine.) Finally, *independent discovery* is not actionable in copyright; absent the copying of another’s work, there can be no liability. We discuss some reasons why this rule makes sense in Chapter 5.

Most of the time, the prevailing plaintiff in a copyright infringement action obtains an injunction – although the U.S. Supreme Court has recently

⁶³ Some limited moral rights protection under the Visual Artists Rights Act and some other bodies of law. See Cotter (1997). Neighboring rights often confer a degree of protection, falling short of a full-blown copyright, upon performers and other entities not covered by copyright law. In the United States, § 1101 of the Copyright Act confers neighboring rights protection upon musical performers only.

⁶⁴ See 17 U.S.C. §§ 109(a), (c). The first-sale doctrine does not apply to computer programs and sound recordings, however, see *id.* § 109(b), because, as one commentator has noted with respect to these works, “the dangers of renter copying are particularly apparent.” Netanel (1996; 300). Moreover, the owner retains a right to prohibit the display of the work by projection of more than one image at a time or to viewers not present at the place where the copy is located. See 17 U.S.C. § 109(c).

⁶⁵ See 17 U.S.C. §§ 108, 110–22. For example, § 114 provides some exceptions to the copyright owner’s exclusive rights in sound recordings in connection with, *inter alia*, webcasting. In a standard edition of the U.S. Copyright Act, this section takes up nearly twenty pages of text and is virtually incomprehensible to nonspecialists.

cautioned that injunctions are not automatic.⁶⁵ In addition, there are a few discrete situations in which copyright law mandates compulsory licensing instead of an injunction.⁶⁶ Under U.S. law the victorious copyright owner is entitled to his “actual damages and any additional profits of the infringer that are attributable to the infringement and are not taken into account in computing the actual damages.”⁶⁷ Normally, this means that the plaintiff is entitled to the larger of either (1) his own lost profits or (2) the defendant’s profits attributable to the infringement. As we shall see, these numbers need not be identical. To recover the defendant’s profits, once the copyright owner presents proof of the infringer’s gross revenue, the burden shifts to the defendant “to prove his or her deductible expenses and the elements of profits attributable to factors other than the copyrighted work.”⁶⁸ In the alternative, and at the election of the copyright owner, the court may award “statutory damages for all infringements involved in the action with respect to any one work . . . in a sum of not less than \$750 or more than \$30,000 as the court considers just.”⁶⁹ In cases of willful infringement, the court may increase statutory damages to a sum of not more than \$150,000; in cases of innocent infringement, the court may reduce them to a sum of not less than \$200.⁷⁰ The court in its discretion also may award the prevailing party costs and attorney’s fees and (arguably) prejudgment interest, but not punitive damages.⁷¹

Like patent law, copyright can be viewed as performing both an incentive and a prospect-like function. The incentive theory suggests that, in the absence of copyright protection, the number of works created and published would be less than optimal due to the ability of others to free-ride upon the efforts of creators and publishers and thereby prevent them from recouping their investments in creation and publication. At the same time, theorists recognize that too strong a system of copyright protection may

⁶⁵ See *New York Times Co. v. Tasini*, 533 U.S. 483, 505 (2001) and *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578 n.10 (1994).

⁶⁶ For examples under U.S. law, see 17 U.S.C. § 104A(d)(3) that authorizes owners of derivative works based on “restored works,” as defined by *id.* § 104A(h)(6), to continue using derivative works upon payment of reasonable compensation; *id.* U.S.C. § 111(c) that provides compulsory licensing for secondary transmissions by cable systems; and *id.* § 114(d)(2), (f) that provides compulsory licensing of copyrights in sound recordings for use in digital transmission subscription services.

⁶⁷ 17 U.S.C. § 504(a).

⁶⁸ *Id.* § 504(b).

⁶⁹ *Id.* § 504(c)(2).

⁷⁰ See *id.*

⁷¹ See *id.* § 505; 3 Nimmer (2002; § 14.02[B], at 14–24 to –28).

: a few
ens
owner
ringer
unt in
tiff is
dant's
members
yright
ifts to
ments
In the
award
spect
000 as
t may
uses of
s than
'costs
nitive

entive
ie ab-
lished
upon
m re-
time,
i may

Music,
ivative
deriva-
ovides
(d)(2),
digital

for the creation of new works that build upon earlier ones due to the presence of transaction costs and other bargaining obstacles that may restrict access to these earlier works.⁷² The prospect theory suggests that according to ownership rights in all of the various uses for any given copyrighted work will maximize social welfare by encouraging the efficient development of markets for those uses.⁷³ Both theories are subject to the same general critiques discussed previously, as well as some new ones.⁷⁴ Moreover, in some cases, the two theories can produce conflicting policy recommendations. Providing the copyright owner with an exclusive right to prepare derivative works, for example, may be difficult to justify on the basis of an incentive theory alone, because in most cases the additional incentive to creativity attributable to this right will be small.⁷⁵ From the standpoint of prospect

theory, see Landes & Posner (1989); Sterk (1996; 1204–5). The administrative costs of the copyright system are probably lower than the corresponding costs of the patent system, however, due to the low level of scrutiny that the Copyright Office employs before registering a work. In addition, monopoly costs are in general probably lower as well, due to the presence of substitutes for many copyrighted works and to a variety of copyright doctrines, such as merger and scenes à faire. Under the merger doctrine, when there are only a small number of ways of expressing a given idea, the idea “merges” with the expression and no copyright may subsist in the expression. The scenes à faire doctrine prohibits copyright protection for standard characters or plot devices or for standard computer programming techniques. See Goldstein (1994; 178–9) (“The logic of property rights dictates their extension into every corner in which people derive enjoyment and value from literary and artistic works. To stop short of these ends would deprive producers of the signals of consumer preference that trigger and direct their investments.”); see also Netanel (1996; 308–36) who critiqued this theory.

Some recent work purports to show, for example, that under certain conditions, content providers may be *better off* if (1) they charge a higher price to the initial purchasers of copyrighted works, and (2) then permit the purchasers to make unlimited copies. See Boldrin & Levine (2002); Watt (2000; 24–70); and cf. Klein et al. (2002). Others argue that in the digital environment, the production and dissemination of works of authorship can be underwritten by content *users'* investment in computer hardware and software. See Ku (2002). Alternatively, one could attempt to reimburse content providers by imposing a tax upon copying equipment and then directing some portion of the tax revenues to content providers rather than by imposing copying restrictions. This approach has made limited headway in the United States but has been used to a greater extent elsewhere. See Watt (2000; 132–4). In addition, when network effects are present, copyright protection can result in very strong rights to control the direction of an industry. See Lemley & McGowan (1998). In these instances, the monopoly costs of protection may be very high. Finally, the cost and efficacy of self-help measures such as encryption may be different in the digital environment than in the nonvirtual world. Whether this phenomenon renders copyright unnecessary, or calls for ever more vigilance to protect some of the values (such as fair use) embodied in traditional copyright law, remains a matter of intense debate.

See Sterk (1996; 1215–17). Sterk recognizes, however, that under some circumstances, the expectation of derivative revenues may be a motivating factor in creating the original work, but he argues that these circumstances are atypical. And there are other possible utilitarian

theory, on the other hand, the adaptation right may seem desirable because it facilitates the copyright owner's ability to efficiently coordinate investment in specific derivative works for which consumers are willing to pay and reduces the probability that overuse will cause the value of a copyrighted work to fall.⁷⁶

TRADEMARKS

Yet another source of intellectual property law, in addition to patents and copyrights, is the law of unfair competition. In one sense, classifying this body of law under the same general heading as patents and copyrights is problematic, because in many respects the rationale for its existence is quite different from the rationales that underlie the other two. Moreover, the term "unfair competition" itself encompasses many bodies of law, including the law of trade secrets, as well as trademarks, false advertising, product disparagement, and the right of publicity. For present purposes, we limit our focus to trademarks, that is, words and other symbols that signify a unique source or sponsor of a product or service.⁷⁷

Under the expansive view that prevails in most industrialized countries today, a trademark can be *any* symbol that identifies a unique product or service. The most obvious examples are words, such as COCA-COLA or MICROSOFT, but trademark rights can subsist in other distinctive symbols as well, such as pictures (the MICHELIN man for tires), numbers (NO. 1 ouzo), and letters (ABC, for a variety of products and services including the television network). Even attributes such as colors (the color pink for fiberglass), fragrances (the scent of plumeria blossoms applied to yarn),

justifications for the adaptation right. In the absence of such a right, the first person to make a particular adaptation may effectively preempt the field, discouraging anyone else from adapting the same work. This outcome is undesirable if the first adaptor is not as talented as another adaptor would have been. Alternatively, perhaps no one would bother to create a resource-intensive derivative work, if others were free to create competing derivative works based upon the same underlying work. See Netanel (1996; 379).

⁷⁶ See 2 Goldstein (2002; § 5.3, at 5:81); Landes & Posner (2002; 13–15); and Landes & Posner (1989; 354–5). See Lemley (1997; 1044–77) where a critique of prospect-theoretical justification for the current scope of adaptation rights is provided. These differences of opinion concerning the appropriate scope of copyright law tend to divide law and economics scholars who write in this field into two camps, which Neil Netanel refers to as the "minimalist" and "neoclassical" schools. Netanel (1996; 309–11).

⁷⁷ Virtually any symbol, including colors, sounds, fragrances, product packaging, and product configuration can serve as trademarks, as long as they are sufficiently distinctive and nonfunctional. See *Wal-Mart Stores, Inc. v. Samara Bros.*, 529 U.S. 205 (2000) and *Qualitex Co. v. Jacobson Prods.*, 514 U.S. 159 (1995).

because
stment
nd re-
d work

ts and
ig this
ghts is
s quite
er, the
uding
oduct
e limit
nify a

ntries
uct or
LA or
mbols
NO. 1
uding
ik for
(yarn),

o make
e from
lented
create a
works

Posner
justifi-
pinion
schol-
nalist”

prod-
ve and
ualitex

sounds (the NBC chimes), and “trade dress” – a term that can refer to product packaging, product shape, and even restaurant decor – can serve as source signifiers, subject to the qualifications that they be distinctive (meaning that consumers are believed to perceive them as source signifiers) and nonfunctional (meaning, roughly, that the attribute does not significantly affect the cost or quality of the article of which it is a part or place nonusing competitors at a significant nonreputation-related disadvantage).⁷⁸

In most countries other than the United States, a person acquires trademark rights through registration, although subsequent use of the mark within a specified period of time is usually necessary to keep the registration in force. In the United States, by contrast, use is generally a prerequisite to protection: the first person to make a lawful, commercial use of a symbol to identify her product or service acquires a trademark right by operation of law, although this right may be enforceable only within the geographic area in which the product or service has been sold or advertised.⁷⁹ Since 1947, however, the United States has eased the common-law rule by permitting the federal registration of marks used in interstate commerce. Registration under the federal Lanham Act creates a presumption of nationwide rights, generally enforceable even in areas outside of the trademark owner’s actual market.⁸⁰ Since 1989, firms also may effectively reserve some marks for a period of time prior to actual use by filing an intent-to-use (ITU) application, although no actual trademark rights spring into existence unless and until the owner begins actual use.⁸¹ Subject to some exceptions – including a first-sale or exhaustion doctrine, similar to the one we have encountered in patent and copyright law⁸² – trademark rights may persist for as long as

some symbols – for example, “fanciful” words such as KODAK – are viewed as being inherently distinctive, meaning that consumer identification of the symbol with a unique source is presumed. Trademark protection springs into existence upon first use (or first filing, in most countries), without the need to prove that consumers actually identify the word with a single source. On the other hand, if a symbol is noninherently distinctive – for example, a descriptive term (TASTY salad dressing) or product design trade dress – the person claiming trademark rights must be prepared to demonstrate that the symbol has acquired distinctiveness (in trademark parlance, “secondary meaning”) – that is, that a substantial portion of the relevant class of consumers has come to identify the symbol with a unique source. Finally, if a word is the generic term for a class of products, it cannot serve as a trademark. Because competitors need to use generic terms to market their products without fear of liability, there can be no SOAP brand soap or CAR brand automobiles. For more detailed discussions of functionality, see Barrett (2004); Thurmon (2004).

See Cotter (1995).

See *id.* at 492 n.24, 536–37.

See 15 U.S.C. § 1051(b).

See Restatement (Third) of Unfair Competition § 24 cmt. b, which stated that “the rights of the trademark owner are exhausted once the owner authorize[s] the initial sale of the

consumers continue to identify the mark with a unique source. To maintain the additional benefits flowing from federal registration, however, the owner must periodically renew her registration.

Ownership of a trademark entails two principal rights. First, and more importantly, the owner has a right to exclude others from the commercial use of a mark that is likely to cause confusion with the owner's mark as to the source or sponsorship of the parties' goods or services.⁸³ For example, if you started selling a soft drink and called it COCA-COLA – or something similar to COCA-COLA – there is a good chance that some nontrivial percentage of consumers coming into contact with your product would mistakenly assume some connection between that product and genuine COCA-COLA products. In this example, your unauthorized use of the words COCA-COLA to sell a soft drink is close to counterfeiting, which itself is a variety of trademark infringement. But trademark infringement goes beyond such obvious examples. Consumers may think, for example, that the words ORLANDO MAGIC on a tee shirt imply that the basketball team has endorsed or authorized the shirt, even though the team's principal business is entertainment and few people would expect the team itself to have manufactured the shirt. In this example, the unauthorized use of a mark to convey a false message of sponsorship is another variety of infringement. Similarly, a small company known as DREAMWERKS that sponsored Star Trek conventions successfully asserted a claim against the Spielberg/Katzenberg/Geffen media giant on a theory of "reverse confusion." Here, the parties were not marketing competitive products or services, and there was little chance that consumers would be confused into thinking that the small company was the source of DREAMWORKS-produced films. The claim was, in fact, the opposite – that consumers would believe the small company was using the name with permission of the giant. Although it is possible that the small firm would benefit from such false association, it is also possible that it could be harmed (through loss of control over its mark, consumer disappointment upon discovering the limited nature of DREAMWERKS's services, and so on); therefore, this possibility is sufficient to state a claim under a theory

product under that trademark"; see also *Champion Spark Plug Co. v. Sanders*, 331 U.S. 125 (1947), that held that the use of the original trademark on reconditioned spark plugs did not infringe, where accompanied by an appropriate disclaimer. As with patents and copyrights, the issue sometimes arises whether the first sale that "counts" for purposes of exhaustion is the first sale within the country in which protection is sought or the first sale anywhere. U.S. law adds to the confusion over this issue by introducing a partially overlapping body of customs law to the equation. But these are issues we do not take up in this work.

⁸³ See 15 U.S.C. §§ 1051, 1072, 1115.

ntain
wner
more
ercial
to the
if you
milar
ige of
sume
prod-
LA to
rade-
vious
NDO
itho-
ment
shirt.
ge of
pany
cess-
giant
eting
mers
ce of
ite –
ame
firm
d be
nent
d so
eory

of "reverse confusion."⁸⁴ Under modern law, the gist of the infringement cause of action is that a *substantial* portion of the likely purchasers of the products or services at issue is *likely* to be confused as to the source or sponsorship of *either* party's goods or services. Courts consider a variety of factors in determining whether confusion is likely.⁸⁵ Significantly, there is no requirement that consumers know the identity of the unique source the mark signifies (how many people know, for example, that COCA-COLA is the ultimate parent company of SPRITE?); or that the products be competitive (e.g., DREAMWERKS); or that any consumers have been *actually* confused; or that the plaintiff prove that the expected confusion will be material to consumers' purchasing decisions. There are, however, a variety of defenses and doctrines that confine the trademark owner's rights in certain discrete cases.⁸⁶

A second right, applicable only to famous, highly distinctive marks,⁸⁷ is the right to prevent trademark dilution. Dilution is the lessening of the capacity of a mark to identify a unique product or service.⁸⁸ For example, when you hear the word KODAK, you immediately think of film, but if the word KODAK appeared (even nonconfusingly) on other products (e.g., KODAK trees, KODAK cola, and so on), it would no longer call to mind a single product or product source. The principal theory behind antidilution laws is that such uses diminish the value of the mark; the benefit of these laws to consumers is more difficult to discern, although there may be some benefit in,

⁸⁴ See *Dreamwerks Production Group, Inc. v. SKG Studio*, 142 F.3d 1127 (9th Cir. 1998).

⁸⁵ Most are variations on the factors set forth in the Restatement (Third) of Unfair Competition § 29. They include, *inter alia*, the similarity of the marks; the inherent and acquired distinctiveness of the plaintiff's mark; the proximity of the plaintiff's and defendant's goods; the degree of purchaser sophistication; the manner in which the goods are marketed; any evidence of actual confusion; and the defendant's intent. See also Cotter (1995; 530).

⁸⁶ For example, the nonconfusing use of another's mark either to describe one's own products or to refer truthfully to the owner's products for purposes of comparative advertising can be a lawful use. Similarly, the plaintiff's mark can be invalidated on a number of grounds, including genericness. Many once-trademarked words have become the generic terms for a class of products. Examples include aspirin, trampoline, yo-yo, and Murphy bed.

⁸⁷ The federal antidilution law limits its protection to marks that are famous and highly distinctive. It lists a variety of factors that are relevant to the issue of whether a mark is famous, and courts have come up with additional factors. See 15 U.S.C. § 1125(c)(1). There is, at present, a split of authority on the issue of whether marks must be not only famous but also inherently distinctive to merit protection under the federal law. Courts also differ as to whether fame must be nationwide and pervasive, or whether geographic or niche fame is sufficient. In addition, some states also have antidilution laws that differ in some respects from the federal act.

⁸⁸ See 15 U.S.C. § 1127 (definition of dilution) and Restatement (Third) of Unfair Competition § 29 cmt. e (1995).

for example, preserving against dilution the value of marks that consumers view as status symbols.⁸⁹ Courts further distinguish dilution by blurring, which is at issue in the KODAK example, from dilution by tarnishment, in which the defendant uses the mark in connection with an unwholesome product (for example, ADULTS-R-US for a pornographic website).⁹⁰ In the United States, federal antidilution protection has existed only since 1996, although about half of the states already had parallel antidilution laws as of that date and have retained these laws. A recent Supreme Court decision holds that the federal act provides a remedy only against *actual*, not likely, dilution,⁹¹ but this limitation may not be as significant as it seems given that the definition of actual dilution as a “lessening of capacity” itself seems to contemplate some degree of potential harm. In any event, most of the state laws that address dilution use a “likelihood of dilution” standard and there is a good chance that the language in the federal act that the Court interpreted as requiring a showing of actual dilution was a result of poor drafting that Congress will eventually repair.

The prevailing plaintiff in a federal trademark infringement action is entitled to injunctive relief and, like her counterpart in a copyright case, also may recover the defendant’s profits attributable to the infringement and any damages sustained by the plaintiff as long as the court avoids double counting.⁹² (The Federal Trademark Dilution Act authorizes courts to

⁸⁹ See Kozinski (1993; 969–70) who argues that, to some extent, modern trademark law protects the value of marks as status symbols.

⁹⁰ See *Toys “R” Us, Inc. v. Akkaoui*, 40 U.S.P.Q.2d 1836 (N.D. Cal. 1996).

⁹¹ See *Moseley v. V Secret Catalogue, Inc.* 537 U.S. 418 (2003).

⁹² See *id.* §§ 1116(a), 1117(a). Certain exceptions to damages liability apply, however, with respect to defendants whose only involvement in an infringement is the printing or advertising of an infringing mark. See *id.* § 1114(2). In cases involving counterfeit marks, however – defined as the use of a “counterfeit of a mark that is registered on the principal register of the United States Patent and Trademark Office for such goods or services sold, offered for sale, or distributed and that is in use, whether or not the person against whom relief is sought knew such mark was so registered,” 15 U.S.C. § 1116(d)(1)(B) – the plaintiff may recover treble his actual damages or the defendant’s profits, whichever is greater, see *id.* § 1117(a), (b), or he may elect to recover statutory damages in the amount of “not less than \$500 or more than \$100,000 per counterfeit mark per type of goods or services sold, offered for sale, or distributed, as the court considers just,” 15 U.S.C. §§ 1117(c)(1). If the court finds that the counterfeiting was willful, it may assess statutory damages of “not more than \$1,000,000 per counterfeit mark per type of goods or services sold, offered for sale, or distributed, as the court considers just.” *Id.* § 1117(c)(2). The court is also authorized to order the seizure, upon *ex parte* application, of goods bearing counterfeit marks, 15 U.S.C. § 1116(d), and in cases involving either infringement or counterfeiting to order the destruction of any goods found to bear an infringing or counterfeiting mark, *id.* § 1118. See also 18 U.S.C. § 2320 regarding imposing criminal penalties for trafficking in counterfeit goods or services.

sumers incurring, intent, in lesome? In the 1996, laws as cession is likely, en that ems to the state here is interpreted ig that ion is t case, ement s dou- irts to urk law r, with or ad- marks, incipal s sold, whom aintiff er, see ot less s sold, (1). If f "not ed for uthor- marks, order *id.* § ng in

to obtain the dilution of famous marks and, in cases of willful dilution only, to apply the same set of damages remedies that are available for trademark infringement.⁹³) Historically, however, courts have been reluctant to award the defendant's profits, unless the infringement implies "some connotation of intent" or a knowing act denoting an intent to infringe or reap the harvest of another's mark and advertising.⁹⁴ As in the copyright context, the plaintiff satisfies her burden of production on this issue by providing evidence of the defendant's sales, at which point the burden shifts to the defendant to prove which costs should be deducted to arrive at the correct profit amount.⁹⁵ With respect to actual damages, the court may award either (1) the plaintiff's lost profits attributable to the infringement, (2) the amount necessary to undertake a corrective advertising campaign, or (3) a reasonable royalty for use of the mark.⁹⁶

The Lanham Act also authorizes the enhancement of damages awards in appropriate cases. First, the court may enhance the amount of actual damages by entering a judgment "according to the circumstances of the case, for any sum above the amount found as actual damages, not exceeding three times such amount."⁹⁷ Second:

(1) if the court shall find that the amount of recovery based on profits is either inadequate or excessive the court may in its discretion enter judgment for such sum as the court shall find to be just, according to the circumstances of the case.⁹⁸

In effect, this latter provision enables the court to "[i]ncrease or decrease an award of profits by any amount if the court finds the profit recovery to be either inadequate or excessive."⁹⁹ In general, however, courts usually enhance damages only when the defendant is found to have willfully infringed the plaintiff's mark. The court may award reasonable attorney's

⁹³ See 15 U.S.C. § 1125(c)(1), (2).

⁹⁴ See 4 McCarthy (2002; § 30:62, at 30–101) and Restatement (Third) of Unfair Competition, § 36 cmts. b, c, § 37 & cmt. e.

⁹⁵ See 15 U.S.C. § 1117(a).

⁹⁶ See 4 McCarthy (2002; §§ 30:79–87) and Restatement (Third) of Unfair Competition § 36, which states that the plaintiff's actual damages may include (1) losses resulting from sales or other revenues lost because of defendant's conduct; (2) sales made by plaintiff at prices that have been reasonably reduced because of such conduct; (3) harm to market reputation of plaintiff's goods, services, business, or trademark; and (4) reasonable expenditures made by plaintiff to prevent, correct, or mitigate confusion.

⁹⁷ 15 U.S.C. § 1115(a).

⁹⁸ *Id.* The statute goes on to state that "[s]uch sum in either of the above circumstances shall constitute compensation and not a penalty." *Id.* The precise meaning of this sentence is unclear. See 4 McCarthy (2002; § 30:91).

⁹⁹ 4 McCarthy (2002; § 30:90, at 30–146).

fees to the prevailing party in exceptional cases; whether it should ordinarily award her prejudgment interest as well or whether such relief should be limited to exceptional cases is an issue on which the U.S. courts currently are divided.¹⁰⁰

Trademarks serve several economic functions. First, they lower search costs by allowing consumers to distinguish between products that differ in quality but that, in the absence of differing brand names, would be difficult or impossible to distinguish at the point of purchase.¹⁰¹ In order for trademarks to fulfill this function, however, the goods or services they identify must be of more or less uniform quality. Therefore, a second function of trademarks is to encourage producers to invest in quality control or, more broadly, the development of the consumer goodwill that trademarks symbolize. Put another way, trademarks encourage firms to invest in creating goodwill by conferring a legal right against some, though not all, free-riding upon that goodwill.¹⁰² A third, more controversial, quality of trademarks that may be more relevant to the antidilution cause of action is the ability of trademarks to serve as vehicles for persuasive advertising, which is a function that may be undermined by another's use of a similar mark (even if that use is unlikely to cause confusion) that threatens to "blur" the distinctive nature of the mark or to "tarnish" its image. A fourth, even more controversial, function of trademarks may be to promote monopolistic competition by encouraging consumers to perceive differences among products that are not, in any meaningful sense, distinct. This argument – really more of a critique than an argument in favor of trademark protection – is no longer very popular, but occasionally surfaces in the critical literature.¹⁰³

PROPERTY RULES AND LIABILITY RULES

One issue of considerable importance to the design of intellectual property laws is whether it is preferable to protect IPRs by means of "property rules"

¹⁰⁰ See 15 U.S.C. § 1115(a); 4 McCarthy (2002; §§ 30:91, 30:93).

¹⁰¹ These would be so-called *experience* goods, which must be consumed in order for the purchaser to evaluate their quality. The rationale would not apply to *inspection* goods, which can be examined prior to consumption and evaluated for quality differences.

¹⁰² The anti-free-riding theory does not explain reverse confusion, however, which as discussed earlier is actionable more because of its potential impact upon the small senior user's ability to control the mark. In a reverse confusion case, the defendant's use threatens to destroy the value of the plaintiff's investment but does not constitute free-riding. For further discussion of the law and economics of reverse confusion, see Feldman (2003).

¹⁰³ See Carter (1990; 768), who discusses, but does not endorse this view.

rdinary
ould be
ntly are

search
iffer in
icult or
marks
nust be
emarks
roadly,
ize. Put
oodwill
g upon
ks that
ility of
inction
hat use
nature
versial,
tion by
hat are
re of a
longer

“property
rules”

for the
n goods,
ces.
as dis-
ll senior
threatens
ing. For
2003).

liability rules.” In a famous article, Guido Calabresi and Douglas Melamed (1972; 1092) distinguished these rules in the following manner:

Entitlement is protected by a property rule to the extent that someone who wishes to remove the entitlement from its holder must buy it from him in a voluntary transaction in which the value of the entitlement is agreed upon by the seller. It is the form of entitlement which gives rise to the least amount of state intervention: once the original entitlement is decided upon, the state does not try to decide its value. It lets each of the parties say how much the entitlement is worth to him, and gives the seller a veto if the buyer does not offer enough. . . .

Whenever someone may destroy the initial entitlement if he is willing to pay an objectively determined value for it, an entitlement is protected by a liability rule. . . . Obviously, liability rules involve an additional stage of state intervention: not only are entitlements protected, but their transfer or destruction is allowed on the basis of a value determined by some organ of the state rather than by the parties themselves.

As our previous discussion shows, the law generally entitles the owner of an IPR to obtain an injunction against the unauthorized use of his patent, copyright, trademark, or trade secret; in so doing, the law encourages the owner and the would-be user to bargain for a transfer of rights at a mutually agreed-upon price. In these respects, IPRs are a paradigmatic example of entitlements protected by property rules, although there are some instances in which the would-be infringer is entitled, as under a liability-rule system, simply to “breach and pay damages” whenever he wishes to use another’s intellectual property.¹⁰⁴

¹⁰⁴ For examples in copyright law, see *supra* note 66. In the law of trade secrets, as we have seen, in exceptional cases a court may permit the defendant to continue using the secret upon payment of a reasonable royalty. In addition, an employee who is deemed to own an invention or other valuable information created during the period of his employment may be required to provide his employer with a “shop right”—an irrevocable, nonexclusive, royalty-free license to use the invention in its own business. See Cotter (1996; 594). There are a few situations in which the U.S. government has required the compulsory licensing of patents, such as as a remedy for patent misuse or other anticompetitive conduct, and in certain other discrete settings. Other countries may impose compulsory licensing for other reasons, such as to combat national health emergencies. See Cotter (2004). The U.S. government also may effect a taking of intellectual property for a public purpose, subject to the Fifth Amendment’s requirement of just compensation. See, e.g., *Jacobs Wind Elec. Co. v. Department of Transportation*, 626 So. 2d 1333 (Fla. 1993). In fact, the federal government licenses many patents to itself, largely for military purposes. See Cotter (1998).

In a provocative article, Nance (1997; 853) challenges the conventional understanding of Calabresi and Melamed’s distinction between property rules and liability rules, arguing that “property, liability, and inalienability rules should be considered prescriptions concerning what people should do, not descriptions of what they can or must do.” In other words, Nance views the distinction as incorporating a normative aspect. On this understanding, a property entitlement means that others *should not* encroach (but not

Many law and economics scholars have argued that it is generally preferable to protect IPRs through the use of property, as opposed to liability, rules. As Robert Merges has explained, in the context of patents:

[A] property rule makes sense . . . because: (1) there are only two parties to the transaction, and they can easily identify each other; (2) the costs of a transaction between the parties are otherwise low; and (3) a court setting the terms of the exchange would have a difficult time doing so quickly and cheaply, given the specialized nature of the assets and the varied and complex business environments in which the assets are deployed. Hence the parties are left to make their own deal.¹⁰⁵

For these reasons, Merges and other scholars contend that compulsory licensing schemes, under which the owner of an intellectual property right is required to license users at some statutorily fixed or judicially fixed rate, are often less efficient than is a system of property-like protection.¹⁰⁶

To say that property rules are generally preferable is not to say that they are necessarily desirable under all circumstances. As we noted earlier, IPRs sometimes are protected by liability rules instead, whereas in other settings alternative compensation schemes (such as taxes on copying equipment) are in place. These alternatives might make sense if, for example, the transaction costs of bargaining are unusually high and the likelihood of voluntary solution to this problem unusually low. This could occur, for example, if one would need to bargain with many rightholders in order to improve upon a technology and other obstacles block the implementation of voluntary solutions to the transaction cost problem, such as patent pools.¹⁰⁷

necessarily that encroachments will be enjoined), whereas a liability rule means that others *may* encroach (but will be required to pay for doing so). Even under Nance's interpretation, however, IPRs are protected (most of the time) by property rules, insofar as the law prescribes that one should not infringe another's IPR. The damages to which one would be entitled for the period of time preceding the entry of an injunction would be, under Nance's terminology, protected under a "remedial compensation rule." See *id.* at 873, 902–5; see also Coleman & Kraus (1986) (similar).

¹⁰⁵ See Merges (1994; 78) and Merges (1996) where an argument is made that compulsory licensing schemes are suboptimal because they are subject to "legislative lock-in."

¹⁰⁶ See, e.g., Adelstein & Peretz (1985); Gordon (1982; 1613); and Kitch (1977; 286–7). But see Ayres & Talley (1995; 1092–4) for an argument that, under some circumstances, a compulsory licensing system will induce the owner and user to reveal their true valuation of the subject property, and therefore may help to overcome bargaining obstacles arising out of strategic behavior. See also Lemley & Volokh (1998), who argue against the routine granting of preliminary injunctions in copyright infringement cases on First Amendment grounds.

¹⁰⁷ But even then, scholars like Merges have argued that it may be preferable to foster the development of private solutions to transaction-cost problems rather than to impose compulsory licensing. On the other hand, Ayres and Talley have argued that, in some settings, liability rules may induce the owner and would-be user to disclose more information

lly prefer-
s liability,

the trans-
on between
ing would
l nature of
the assets

ulsory li-
right is
l rate, are

that they
lier, IPRs
r settings
(ent) are
transac-
voluntary
ample, if
improve-
n of vol-
pools.¹⁰⁷

that others
s interpre-
far as the
which one
would be,
¹⁰⁷ See *id.* at

ompulsory
n.”
36–7). But
instances, a
valuation
es arising
he routine
endment

foster the
o impose
some set-
formation

Alternatively, there may be cases in which policymakers opt for a liability rule in order to implement a particular vision of social justice. An example might be the recent controversy over the use of compulsory patent licensing to enable greater access to AIDS drugs and other essential medicines in developing countries. Some scholars also are concerned about the overuse of injunctions in copyright and trade secrets cases to the extent that these applications can affect freedom of speech or other constitutional rights.¹⁰⁸ We do not resolve these controversies here. We think it is fair to state, however, that intellectual property law in general prefers property rules to liability rules and that – if the premises upon which the intellectual property system is based are sound – a rebuttable presumption in favor of property rules probably makes more sense than would a presumption favoring liability rules. Deviations from the norm may be justified in some cases, but we will assume that the norm is a property rule.

concerning their subjective valuation of property. See Ayres & Talley (1995; 1092–4). The applicability of this insight to IPRs outside the bilateral monopoly setting has been challenged, however. See Merges (1996; 1304–5); see also Burk (1999; 142).

¹⁰⁷ See Lemley & Volokh (1998).