Free-market economics, the sort celebrated by textbooks and politicians, often seems like pretty straightforward stuff. The virtues of competition, the vices of monopoly – you know the riff. But look closely and you’ll spot plenty of loose ends, none of which is looser (or more important) than the way we treat the intangible “intellectual property” embodied in technology. Hence the significance of Innovation and Its Discontents: How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do About It.* The subject can be tough going. Happily, authors Adam Jaffe of Brandeis University and Josh Lerner of the Harvard Business School, write almost as well as they think. Heck, they even manage to inject some humor: where else are you going to learn about the patent on a wrist- (paw-?) watch for dogs that rushes through the hours at seven times normal speed? — Peter Passell
In 2001, Albie’s Foods, a small grocery and caterer, received a letter from the law firm of jam and jelly giant J.M. Smuckers. It accused Albie’s, which markets pastries and sandwiches in northern Michigan, of violating Smuckers’ intellectual-property rights by selling peanut butter and jelly sandwiches without the crusts.

In particular, Smuckers claimed that Albie’s had infringed Smuckers’ recently granted U.S. Patent 6,004,596 covering a “Sealed Crustless Sandwich.” Indeed, as the patent abstract noted, Smuckers’ had been granted broad protection:

The sandwich includes a lower bread portion, an upper bread portion, an upper filling and a lower filling between the lower and upper bread portions, a center filling sealed between the upper and lower fillings, and a crimped edge along an outer perimeter of the bread portions for sealing the fillings there between. The upper and lower fillings are preferably comprised of peanut butter and the center filling is comprised of at least jelly.

Probably surprising the jam magnates, Albie’s decided to defend itself in federal court. Albie’s lawyers noted that the “pasty” – a pocket sandwich with crimped edges and no crust – had been popular fare in northern Michigan since the immigration of copper and iron miners from Cornwall, England in the 19th century.

A battle in federal court over peanut butter and jelly sandwiches may seem merely funny, and a little pathetic. But it is symptomatic of the larger and more profound problems with the patent system.

A PATENT PRIMER

In the United States, patents are issued by the Department of Commerce’s Patent and Trademark Office (PTO). A patent describes an invention, and lists one or more claims that specify what the invention does that has never been done before. The holder of a patent has the legal right to prevent anyone else from making, using or selling an object or device that incorporates any feature covered by the specified claims. This right operates within
the United States, but also bans imports whose manufacture violates the patent.

**The Process of Patent “Examination”**

To get a U.S. patent, the inventor files an application with the PTO. A patent office examiner must then determine whether the invention meets the standards for patentability. First, the subject matter of the invention must be suitable. U.S. law permits the granting of a patent on the following subjects:

- A process, such as a new approach to brewing beer or to depositing circuits on silicon.
- A machine, such as a new machine tool or automobile carburetor.
- An article of manufacture, such as a kit to identify an infectious disease.
- A composition of matter, such as a novel type of concrete or a new molecule.
- New and useful improvements of the above.
- Any distinct and new variety of plant that is asexually reproduced.
- Any new, original and ornamental design for an article of manufacture.

But that is not the end of the story. The patent application must also pass three other tests:

**Utility.** Does the invention really do anything, and if so, does it solve the problem it set out to address? In practice, the requirement that an invention must be useful is not of great importance because it is easy to make the case that almost anything is potentially useful – for example, U.S. Patent No. 5,023,850 covers a wristwatch for dogs (paw-watch?) that moves at seven times the speed of an ordinary clock.

**Novelty.** Is the claimed invention really original?

**Non-obviousness.** Even if new, would the claimed invention have been obvious to one skilled in the art at the time of the invention? The non-obviousness requirement is a supplement to the novelty requirement, preventing one from getting a patent on something that hasn’t been done before, but easily could have been. Thus, if a skilled practitioner could easily have modified some existing invention to arrive at the new invention, it is not patentable.

The examiner’s job is, in effect, to compare the invention to the “prior art” – everything that was known before the time of the application. To get a patent, the invention must pass the tests of novelty and non-obviousness in relation to the prior art. A patent application begins with a general description of the invention and what it does. But the legal guts of the patent are its “claims.” Patent claims are the legal characterization of what is and is not covered by the patent; hence they are worded in a very specific, legalistic way. A theoretical patent on a mousetrap might have an abstract saying something like:

A device for catching mice or other mammalian pests not exceeding 100 grams, wherein a levered bar is attached to a spring or similar device, while being restrained by a hook, clasp or similar device, which is in turn attached to a small plate containing pâté de foie gras. The connection between the pâté-containing plate and the hook or clasp is such that any disturbance of the pâté releases the levered bar. The levered bar is positioned in such a way that upon release it falls where the head or neck of a rodent is likely to be located if said rodent is attempting to eat said pâté. The spring or similar device causes this movement of the bar to occur with sufficient force that said rodent is killed, incapacitated or restrained in such a way as to prevent its escape.

Notice that this abstract contains words...
or phrases that make the meaning as broad as possible, such as “or other ... pests” and “or similar device.” Such broad language is intended to make the patent cover as many different devices as possible, giving the holder of the patent maximum power to restrict competition. But notice also that the patent contains language that limits its coverage, such as “mammalian” and “not exceeding 100 grams,” and is quite specific as to the bait used. Such restrictions distinguish the invention from earlier ones, in order to establish novelty and non-obviousness.

If, for example, there was a previous device of similar design that was used to catch reptiles, or mammals exceeding 100 grams, the patent as drafted distinguishes the new invention from that prior art. It would still be up to the examiner, however, to determine whether adapting some previous reptile trap to mice was obvious or not. Similarly, while it is likely that no one has ever before proposed using pâté as mouse bait, this distinction could justify a patent only if it were judged that using such high-class bait instead of cheese or peanut butter constitutes an inventive leap that would not be obvious to a skilled mousetrap designer.

If the examiner concludes that the claims are not novel, or are obvious, the applicant has the opportunity to redraft the claims to make them more restrictive. The process of patent examination is therefore, to a large extent, one of negotiation between the applicant (typically, the applicant’s lawyers) and the examiner. The applicant generally wants claim language that is as broad as possible. The examiner may, however, insist on restrictions to distinguish the invention from the prior art. And because the applicant has unlimited opportunities to amend the application to satisfy the examiner, a large fraction of all original patent applications are now ultimately granted.

A key feature of this process is that it involves only the applicant and the PTO – others are not allowed to participate. Indeed, until recently, the entire application process was secret. Even the existence of the application was kept secret until a patent was granted, and was kept secret forever if a patent was not granted. But under a change made in 1999, most patent applications in the United States are now published 18 months after filing.

**So You Got Yourself a Patent ...**

If the PTO grants a patent, the patentee has a legal monopoly for 20 years from the date of application. While the applicant must be a
human being, the inventor can assign the patent to a company. In fact, it is general practice for firms to require their employees to assign their patents to them, and, typically, the firm’s lawyers handle the application.

Like any other property, a patent can be bought and sold, left in a will, given as a wedding present, or left in the attic and forgotten. And, just like a house or a piece of a land, a patentee can rent the patent to someone else by granting a license to use the technology. The license agreement may or may not require royalty payments. And royalties can be in any form the parties agree to – anything from a lump-sum payment to a percentage of profits.

The big difference between intellectual property and tangible property, however, is that the former can be “rented” to multiple people at the same time, with or without restrictions on how each licensee uses the technology. Of course, in many cases, the patentee uses the technology in its own business, and relies on the patent to prevent competitors from using the technology.

When a patentee learns that someone is infringing its patent, it will typically write a letter demanding that the alleged infringing activity stop. The letter may offer to license the patent in return for a royalty. The recipient of such a letter essentially has three choices. He or she can agree to pay for a license, or stop doing whatever is alleged, or continue as before and wait for the patentee’s next move.

In many cases, the parties try to resolve the dispute through negotiation, which may lead to a license agreement. Or the patentee can try to stop the alleged infringement by initiating litigation in federal court to enforce the patent. If the patentee can prove that someone is infringing the patent, he is entitled to an injunction ordering the infringing activity to cease. He is also entitled to compensation for the infringement.

The defendant in an infringement suit will typically counter-sue, claiming that the patent is invalid. There are a variety of grounds for arguing invalidity, including fraud or some other malfeasance. But the most straightforward basis is that the examiner made a mistake and that the invention is not novel or is obvious. Thus, to prevail, the patentee needs the court to decide that the patent is valid and is being infringed. The alleged infringer can prevail via either a finding of non-infringement or a finding that the patent is not valid.

**A Mess of Peanut Butter and Jelly**

Let’s now follow the PB&J patent through this process. In December 1997, Len Kretchman of Fergus Falls, Minn. and David Geske of Fargo, N.D. filed an application for a patent on a “Sealed Crustless Sandwich.” The primary examiner was Lien Tran. Two years later, the patent was granted, containing 10 claims. The first claim reads as follows:

A sealed crustless sandwich, comprising:

- a first bread layer having a first perimeter surface coplanar to a contact surface;
- at least one filling of an edible food juxtaposed to said contact surface;
- a second bread layer juxtaposed to said at least one filling opposite of said first bread layer, wherein said second bread layer includes a second perimeter surface similar to said first perimeter surface;
- a crimped edge directly between said first perimeter surface and said second perimeter surface for sealing said at least one filling between said first bread layer and said second bread layer;
wherein a crust portion of said first bread layer and said second bread layer has been removed.

In simple English, the examiner’s decision to allow these claims amounts to a determination that surrounding jam with peanut butter so the bread won’t get soggy is a new idea, and one that was not previously obvious to skilled sandwich makers.

In the process of examining this application, relevant prior art was identified in the form of seven previous patents issued between 1963 and 1998, along with a 1994 book entitled Fifty Great Sandwiches. Now, it may strike the reader that one needn’t look to earlier patents for the evidence that cutting the crusts off a sandwich, surrounding the jam with peanut butter, or even crimping the edges, were probably not new ideas. Indeed, for some technologies, such as software, a big problem with the examination process is that the prior art exists in forms other than patents, and patent examiners are not good at finding non-patented prior art.

In this case, however, there was evidence undercutting the validity of the patent within the files of the PTO. In particular, U.S. Patent No. 2,463,439, issued in 1949, had been for a device creating very similar kinds of sandwiches. Not only did a family of Cincinnati restaurateurs patent this discovery, it sold a device based on this discovery. On eBay, for instance, one occasionally sees a “Toastite” from the 1950s that produces a “sealed crustless sandwich” akin to that described in the old (and new!) patent.

But as the applicants had included neither the earlier patent nor any information about pasties or Toastites in their patent application, the patent examiner did not find this particular prior art, or else somehow judged it irrelevant. So the patent was granted, and Albie’s got the cease-and-desist letter. Many small firms in their position would have simply dropped the pasty from their culinary portfolio, or agreed to pay a royalty if Smucker’s demand was small enough. But Albie’s chose to fight. Eventually, as in so many of these disputes, the two parties reached a private settlement.

What’s Gone Wrong?
The granting of patents despite clear evidence of invalidity in the form of prior art has become all too common. Staying in the bread area, there is U.S. Patent No. 6,080,436 (“Bread Refreshing Method”), which is an “invention concerned with the process and apparatus for refreshing bread products, particularly open face items such as sliced rolls, buns, muffins, and the like … via exposure to high heat” – what most people would call toasting. Anyone who has recently browned a slightly stale hot dog bun over a barbeque has probably infringed this award!

How could the patent office have missed these cases, where there is so much prior art and where the technology is so basic? Such patents – as well as peers such as “Method of Exercising a Cat” and “Method of Swinging on a Swing”* – are just the amusing tip of an ominous iceberg. As a result of legal and administrative changes made between 1982 and 1990, the PTO has become so overtaxed and its incentives have become so skewed toward granting patents that the tests for novelty and non-obviousness have become largely non-operative. Simultaneously, changes in the court system have made patents much

*The PTO ordered the patent re-examined after it attracted considerable press attention for having been awarded to a five-year-old. (His father was the patent attorney.)

pressure to preclude the escape of the filling from
more powerful legal weapons: a patentee is more likely to win an infringement suit against a broader array of possible infringers than was the case before.

In other words, in the space of less than a decade, we converted the weapon that a patent represents from a pocketknife to a bazooka, and then started handing out the bazookas to pretty much anyone who asked. The result has been a dangerous and expensive arms race, which now undermines rather than fosters the crucial process of technological innovation.

The key to getting this system back on track is to restructure the incentives of all of the parties – the PTO, potential applicants, other inventors and patentees – to reduce the flow of applications, improve the rigor of examination, and reduce the incentive to use patent litigation as a competitive weapon. But before making that case, we need to explore what it is that patents are supposed to accomplish – and why it is devilishly tricky to get just the right level of rigor in the patent system.

**WHY DO WE NEED A PATENT SYSTEM TO BEGIN WITH?**

Patents are government-sanctioned monopolies. Monopolies are bad. Why not abolish the patent system altogether?

**Two Stories**

Qualcomm, a communications-technology company founded in 1985, pioneered a technology known as CDMA that allows cell phones to use the radio spectrum more efficiently. CDMA is covered by a patent requested by Qualcomm in 1990 and granted in 1992. Qualcomm now has revenues of approximately $3 billion, with about $800 million
coming from royalties earned on its patent licenses and the remainder from its own product sales.

To get to this point, Qualcomm has had to compete with large, established companies like Motorola and Nokia. It is unlikely that this success could have been achieved if the CDMA technology had not been protected by patent.

Biogen, founded in 1978 by a group of scientists, including two who eventually won Nobel Prizes, was among the first companies to use the new techniques of genetic engineering to develop pharmaceuticals. It raised millions of dollars to fund the development of new products, including alpha- and beta-interferon and a vaccine for hepatitis B. Neither its investors nor its industry partners would have been willing to put resources into these products if patents had not protected them. Indeed, for the first 15 years, Biogen never manufactured or sold any tangible product, generating revenues solely by licensing or selling its research to other companies. Such specialization would not be possible without patents.

The Simple Case for Patents

The standard economic explanation of the societal function of patents has three basic parts. First, the development and commercialization of new technologies – technological innovation – creates broad societal benefits. It is thus in our collective interest to create legal institutions that foster technological innovation.

Second, though a spark of creative genius may ignite an innovative flame, it typically takes years of research and development to nurture the fire into a commercially viable blaze. And that process costs money – often a lot of money.

If technological innovation is socially desirable but expensive, society needs to have institutions that direct time and money into the process. One approach might be to have the government use tax money to develop new technologies. In fact, the U.S. gov-
ernment does do a lot of R&D, particularly in areas like defense, space and the environment that are themselves important areas of government responsibility. Yet, while government is good at many things, entrepreneurial initiative is not one of them. We thus need institutions that create financial incentives for individuals and firms to invest their own money in the process.

Under a patent system, this incentive takes the form of a zone of economic exclusivity for the innovator. A more direct approach might be simply to reward the innovator with money: In the 17th and 18th centuries, the early patent system in Great Britain coexisted with the use of “prizes” to reward people who solved particular technological problems. But, while prizes may be an effective mechanism for drawing forth a specific, desired technology, they are not as effective at bringing forth innovation in general.

First, it would be expensive to hand out enough prizes to reward the gamut of industrial innovation. Further, it would be hard to figure out how big a prize to give to each innovation. For instance, the British government paid a full £50,000 (several millions of today’s dollars) to John Palmer, the inventor of a new way of organizing the mail – far more than that paid to Edward Jenner for the smallpox vaccine that saved millions of lives. The reward to patents, by contrast, is by its nature proportional to the size of the discovery: the exclusive right to a modest discovery is unlikely to be worth very much, while the exclusive right to an important new technology is usually very valuable.

Building a Better Mousetrap
Suppose you build a better mousetrap. The world beats a path to your door, and everyone coming up that path pays you handsomely for dispatching their rodents with such efficiency. So far, so good, except that all of the other mousetrap manufacturers are likely to buy a couple of your new traps, and figure out how they work. If they can, they will copy your design and sell their own versions.

While there were, in fact, 34 patents issued in the past 25 years in the United States for mousetrap-related inventions (including “Rubber-band Powered Mousetrap”), it is not a particularly important economic sector. But the same economic dynamic is at work in more important industrial sectors – for example, with prescription pharmaceuticals. The Food and Drug Administration must approve drugs sold in the United States. The FDA requires proof, in the form of expensive clinical studies, that a drug is both safe and effective.

This means that after a chemical compound has been discovered and its initial use for health purposes identified, someone (usually a large drug company) must spend hundreds of millions of dollars conducting tests. Once the drug is on the market, however, anyone with a decent chemistry lab can figure out what it’s made of, and in most cases could manufacture and sell it.

Competition would then drive down the price of the drug. Lower prices might seem like a good thing. But if competition quickly reduces the price of a new drug, the company that invented it would not earn significant profits from selling it. And this is exactly what one would expect to happen in a world without patents.

Now, put yourself in the lab coat of a pharmaceutical R&D director. If a promising drug turns out to be a dud, you lose a couple hundred million dollars. If it turns out to be successful, your competitors jump in and sell it, too, so you wouldn’t really make much money. This is not a game that your board of directors is going to let you play for very long.
And if no one plays this game, no new drugs get developed.

Thus, patents make new drugs expensive, which is bad. But if they weren’t expensive, nobody could justify the large cost of developing them. And expensive new drugs are better than no new drugs. This is the tradeoff at the heart of the patent system.

**Controlling the Risks Inherent in Innovation**

Most investments in new products and processes fail. Thus, for the overall “game” of investing in new technology to be worthwhile, the successes must earn enough to cover not only their own costs and a reasonable return, but also the costs and a reasonable return for all of the failures.

Studies of pharmaceutical R&D suggest that if you screen 5,000-10,000 compounds for possible clinical use, on average 250 will show enough promise to be put into preclinical testing. Of these 250, five will show sufficient promise to enter clinical testing; the rest are simply abandoned. Clinical testing is where the really big bucks get spent, typically in excess of half a billion dollars per drug. Of these five drugs subjected to expensive clinical testing, on average only one will be approved by the FDA. And of course, FDA approval does not ensure large profits. Many approved products have small markets.

If we combine the risks of failure at the clinical-trial and market stages, 80 percent of these expensive testing efforts are complete losers, never leading to an FDA-approved product. About 14 percent lead to an FDA-approved product, but do not earn enough to recoup their own development costs. Only 6 percent earn enough to recoup their own investment costs — and, of course, this 6 percent must also earn sufficient profits to pay for all of the losses on the other 94 percent of compounds tested!

Add to this the problem that neither the managers of firms nor investors like risk. All else being equal, investments that are risky are less likely to be undertaken. Investment in new technology is therefore handicapped by its riskiness, when compared with other forms of investment (for instance, expenditures on marketing an existing brand). Furthermore, when a business builds a new factory or buys new equipment, it doesn’t normally worry that its competitors will steal the assets. By contrast, intangible R&D is relatively easy to steal. This is where the patent system comes in.

**THE SIMPLE CASE IS TOO SIMPLE**

The above explanation of the desirability of patents contains a lot of truth, but it also leaves a lot out:

- Companies have other methods for preventing imitation.
- Inventions don’t occur in isolation. If one firm gets a patent, it can stifle inventions that other firms might otherwise undertake, thereby clouding the patent’s overall effect on technological progress.
- Some patentees use patents more like grenades than security fences, threatening others’ property rather than just defending their own.

**Other Means of Protecting New Products and Processes**

There is a safe somewhere in Atlanta in which (it is said) the formula for Coca-Cola resides. This formula is not patented, but because the drink is a mixture of complex natural substances it is not possible to duplicate Coke
simply by analyzing the stuff in a spectrometer. Similarly, even if there were no patents on components of Microsoft Windows, it would be virtually impossible for a would-be imitator to create an exact duplicate of this fantastically complex PC operating system.

As a legal matter, trade secrecy is a distinct form of protection for intellectual property. If a company makes appropriate efforts to prevent the disclosure of information that is valuable to its business, competitors who get access can be sued if they attempt to use it commercially. Secrecy can be particularly valuable in protecting new processes rather than new products.

Another way that firms can protect innovations without patents is through a first-mover advantage. In the mouse trap imitation story told above, once the competitors come up with their copies, they are able to take away many of your customers. But sometimes being first is enough to confer a lasting competitive advantage. Once consumers have found a mouse trap that really works, they may be reluctant to fool around with an imitator.

None of this implies that patents are unimportant. But it does mean that their importance in maintaining the flow of new technologies varies across different industries.

Cumulative and Overlapping Innovation

The firm selling a technology has an incentive to improve it. But in many cases, the investor may not be the one in the best position to make improvements. So how do we create broad incentives for people to invest in cumulative innovation?

The obvious answer is to grant patents on improvements, and such patents are indeed allowed if the improvement embodies an idea
not covered by the patent on the underlying technology. This seems fair, but in practice it is tricky to implement. To make the patent on the original invention useful, its owner must be given some latitude to modify the invention and still have it covered by the original patent. But if the latitude is too wide, many improvements are likely to fall under the original patent. So there is a tradeoff: granting broad patent protection creates the maximum incentive for original inventions, but it may actually discourage improvements.

A classic example is the electric light bulb. Thomas Edison was granted the basic patent on incandescent lighting in 1880. Many companies offered competing products, a number of which improved on the original design; by no coincidence, the cost of bulbs trended steadily downward. Then, in 1891, Edison General Electric Company won an infringement suit against competitor United States Electric Lighting, and subsequently won injunctions against a number of other competitors. Improvements slowed until the expiration of the patent allowed the competitors to resume their efforts.

Now surely, Edison’s invention was about as novel as they get. But acknowledging the legitimacy of Edison’s patent and his efforts to enforce it is not inconsistent with recognizing that the monopoly thereby created inhibited subsequent improvements.

In principle, follow-on inventors ought to be able to negotiate an agreement with the owner of the original patent that allows improvements to be implemented. In practice, however, such agreements are often difficult to work out. After the Wright brothers patented their aircraft stabilization and steering system, many others were eager to design new aircraft. But the Wright brothers refused to license their patent for a critical component of aircraft to anyone, and engaged in protracted litigation with a number of designers. With the entry of the United States into World War I, Washington pushed the major aircraft manufacturers, including the Wrights’ firm, to license their pat-
ents as a package to accelerate the manufacture of planes and the development of new designs. The rapid development of numerous aircraft concepts in years following the establishment of this “patent pool” suggests that the pioneering patent (combined with the unwillingness of the inventors to cooperate with their technological followers) retarded innovation.

Thus, there is an inherent tension between providing strong patent rights to encourage breakthrough innovations and the inhibition that those strong protections may create for subsequent improvements. A related problem is created by the reality that firms are often innovating more or less in parallel. In principle, each should be entitled to a patent only on those aspects of the creation that are unique and truly new. In practice, this is very hard to do. What is likely is that each will be granted a patent that describes its invention in a way that leaves considerable ambiguity as to whether the inventions of competitors are covered. In such cases, the patent system may well inhibit rather than encourage innovation.

**Patents Go on the Offensive**

Almost all formal disputes involving patents are litigated in the federal judicial system. Prior to 1982, appeals of patent cases were heard in the appellate courts of the various circuits. The circuits differed considerably in their interpretation of patent law, with some being more than twice as likely to uphold patent claims as others. These differences persisted because the Supreme Court, which normally steps in to ensure national legal uniformity, rarely stooped to hear patent-related cases.

The result was widespread “forum shopping.” Patent applicants would crowd the hallway in the patent office, where the list of patent awards was distributed at noon each Tuesday. Upon discovering that their patents were issued, they would rush to the pay phones to instruct their lawyers to file suits against alleged infringers of the newly minted patents in patent-friendly district courts, such as Kansas City. Meanwhile, representatives of firms who might be accused of infringing the newly issued patents would be racing to the phone bank as well, ordering their lawyers to file lawsuits seeking to have the patents declared invalid in a district known to be skeptical of patents – for example, San Francisco. Such dueling lawsuits would usually be combined into a single action, and heard in the district court in which the earliest filing was made. Often the fate of the case would turn on which lawyer got an earlier date-time stamp on his filing documents.

In 1982, Congress addressed this problem, establishing a centralized appellate court for patent cases called the Court of Appeals for the Federal Circuit (CAFC). The change was presented in the Congressional hearings as a benign one, bringing consistency to the chaotic world of patent litigation. But it was clear that the initial advocates of stronger patent protection hoped that the new court would come down squarely on the side of patent holders.

And this is precisely what happened. The share of cases in which a district court finding of patent infringement was upheld increased, as did the share of cases reversing earlier findings that patent holders were not entitled to damages. Likewise, the CAFC greatly expanded patent holders’ rights along a number of other dimensions – for example, making it easier to shut down a rival’s business even before a patent was proven valid and to extract significantly greater damages from infringers.

This tilt has been exacerbated by changes in the behavior of inventors and of the patent
office, which have led to a dramatic increase in both the number of patent applications and the fraction of applications that are successful. Decisions of the CAFC encouraged more patent applications for three distinct reasons. First, the CAFC made it clear that the realm of patentable subject matter included technologies, like software, business methods and certain kinds of biotechnology, that hitherto were believed by many to be unpatentable. Second, the new court issued rulings on the standards for novelty and non-obviousness that made it easier for applicants to qualify for a patent. Finally, the improved enforceability of granted patents encouraged patent applications by making the patent right more valuable. As a result, the pace of applications began to increase shortly after the creation of the CAFC.

Just as the tide of patent applications began to rise, Congress intervened once again to modify the patent system. In the early 1990s, it converted the patent office from an agency funded by tax revenues to one funded by fees. In fact, the PTO has become a profit center for the government, collecting more than it costs to run the agency.

Again, a seemingly modest administrative change had immodest consequences. The PTO increasingly views its mission as serving patent applicants. Furthermore, the new orientation created strong incentives for the PTO to process applications as quickly as possible, and at the lowest possible cost. As a result, there has been a decline in the rigor with which the standards of novelty and non-obviousness are applied in reviewing patent applications. This, in turn, has encouraged more people to apply for dubious patents.

THE PATENT EXPLOSION
The number of patents granted increased at less than 1 percent annually from 1930 through 1982, the year the CAFC was created. From 1983 to 2002, that rate rose to 5.7 percent. If this increase reflected an explosion in inventiveness, it would be cause for celebration. But clearly a substantial fraction were of dubious merit – a reality confirmed by international comparisons, which show that the number of U.S.-origin inventions with worldwide significance grew in the 1990s at a rate less than half that of U.S. patent office grants.

Much of the problem stems from the organization of the patent office itself. Chronically strained for resources, officials have struggled to find qualified examiners, particularly in the areas of software, financial methods and biotechnology, where it had not previously had much expertise. Examiners of financial patents, for example, often have as little as a dozen hours to assess whether a patent application is truly novel.

Moreover, retaining the few examiners skilled in the new technologies has been difficult. Companies have been eager to hire these examiners, who are valuable not only for their insight into the PTO examination procedure in the new technology, but also for their knowledge of what other patent applications are in process but not yet awarded. Needless to say, corporations and law firms can offer many times the $40,000 starting salaries paid to examiners.

THE PATENT LITIGATION EXPLOSION
The proliferation of patents on previously existing technologies would sow confusion and legal uncertainty under the best of circumstances, but it has occurred just as the CAFC has been making it easier to enforce the rights it conveys.

The number of patent lawsuits was roughly constant over the 1960s and 1970s, but began to rise with the increase in patent awards in
the 1980s and ballooned in the 1990s.

The pernicious consequences can be seen in two sorts of competitive interactions. In one scenario, an established firm—frequently one whose competitive position and innovative activity are declining—realizes it has a valuable stockpile of issued patents. This firm then approaches rivals, demanding that they take out licenses on its patents. In many cases they target smaller firms, which do not have the money to engage in protracted litigation.

Even if the target firm believes that it is not infringing, it may pay to settle rather than fight. For, along with the direct costs of litigation, there are substantial indirect costs associated with patent litigation. The pre-trial proceedings and trial are likely to require the alleged infringer to produce extensive documentation, diverting employees from more productive tasks as well as creating unfavorable publicity. Its officers and directors may also be held individually liable, or be targeted in shareholder lawsuits if the firm's stock price drops.

For numerous large companies—most notoriously, IBM and Texas Instruments—patent enforcement has become a line of business in its own right. Thus, in recent years Texas Instruments has netted close to $1 billion annually from patent licenses and settlements resulting from its general counsel's aggressive enforcement policy. In some years, enforcement revenues have actually exceeded net income from selling products.

In addition to being forced to pay royalties, small firms may reduce or alter their investment in R&D. Smaller firms tend to shy away from research in areas where large firms have established patent portfolios. Thus, these types of enforcement activities by large firms
may have the effect of suppressing innovation by younger, more vibrant concerns.

A second worrisome development has been the emergence of individual inventors seeking to hold up established firms. In many cases, these individuals have received a patent of dubious validity, often with overly broad claims. Yet established players have often chosen to settle such disputes, not wishing to risk the uncertainty associated with submitting a complex piece of intellectual property to trial.

In many cases, inventors demand a jury trial and then present themselves as Davids fighting corporate Goliaths. They may also shop for jurisdictions that are unsympathetic to the defendant. For instance, Jerome Lemelson, who claimed to have invented bar-coding technology, filed suits against Japanese and Korean firms in the Southern District of Texas. Similarly, inventors frequently threaten corporations with preliminary injunctions, which can stop defendants from using the patented technology even before trial. While an established business might be reluctant to ask for such a drastic measure, lest the other party seek a similar ban against itself, individual inventors often feel no such compunction.

**THE SCALE OF THE PROBLEM**

One lesson from the dot-com debacle is that many firms without a sustainable competitive advantage — think of the nine online pet-supply retailers that dueled for customers in the late 1990s — are unlikely to survive. Investors are looking for real yardsticks to evaluate competitive strength. Thus patents, trademarks and other forms of intellectual property represent a currency that is used increasingly to demonstrate to financial markets, suppliers and customers that a firm is a strong player and can be expected to achieve a dominant position.

Though other countries have not imitated the changes in patent procedures in the United States, they nonetheless face a number of the same issues. Proposals to issue software patents have stirred angry debates in Europe. Programmers championing Linux, an “open source” operating system for which developers must eschew intellectual property protection, have led the opposition.

The question of what is patentable is important not only to the health of the economy, but also to the health of our citizens. For instance, advocates have questioned whether the broad coverage that Utah-based Myriad Genetics enjoys on its breast cancer gene patents is slowing research on this disease. Myriad played a key role in identifying the two genes that can trigger breast cancer by exploit-
ing the voluminous genealogical records of the Mormon Church to find critical genetic markers.

For its efforts, Myriad received two patents on diagnostic tests and treatments. It then entered into licensing agreements with leading medical schools, universities and hospitals, giving them the right to research issues related to breast cancer. But these licenses are quite restrictive. Since the first patent was awarded in December 1997, a number of researchers have been forced to abandon their efforts due to these licensing terms.

Even the pro-business Bush administration has raised questions about Bayer’s exploitation of its patent on the antibiotic Cipro. At the time of the October 2001 anthrax scare, advocates pointed out that Bayer was charging the government $1.89 per dose for the antibiotic of choice against this deadly pathogen. While this was less than the drugstore price for Cipro, companies in India were selling generic versions of the drug for less than 20 cents each.

Ultimately, Bayer cut the price it charged Washington by half – but not before the government initiated plans to license the patent to generic drug manufacturers.

WHY IS REFORM SO HARD?

First, the issues are complex, and simplistic claims frequently cloud these discussions. For instance, it is frequently argued that “stronger” patents are beneficial for innovation, and virtually any change to the status quo is characterized as “weakening” the patent system. The lack of ready dialogue between economists and lawyers about these issues has limited the extent to which the discussion can be raised to the more appropriate – but harder to convey – level of how to maximize the patent system’s effectiveness in stimulating innovation.

Second, the people with the greatest economic stake in retaining a litigious and complex patent system – the patent bar – have proven a very powerful lobby. Technology-intensive firms have not mounted an effective counter-campaign around these issues, perhaps because the companies most adversely affected are small, capital-constrained firms that do not have the resources for major lobbying efforts. And even for larger firms, the adverse consequences of a malfunctioning patent system are diffuse and indirect, which works strongly against the emergence of a consistent voice on this subject.

Finally, most of the harm is borne by consumers, for whom the adverse consequences are even less direct and hard to detect. No one sees that products are more expensive because of the cost of litigation and patent royalties, and no one knows about products whose introductions were delayed or cancelled because of patent woes.

So the train is out of control, even if it has not yet jumped the tracks. While the patent system plays a vital role in creating and maintaining incentives for innovation, it is becoming ever more expensive to operate (and fight about). Worse, its pathologies increase the uncertainty associated with investment in innovation, and thus undermine the very incentives it is designed to create.

It is time to recognize that the accidental combination of strengthening the legal value of patents while reducing the rigor of patent examination has damaged the system. Since the source of the mess is the combination of easier success in the courts and at the PTO, our proposed recalibration addresses both of these venues.
Goals and Objectives
While different analysts of the patent landscape have emphasized different aspects of the problem, there is general agreement on broad goals for reform of the patent system:

**Improve Patent Quality.** People are getting patents for inventions that are not new or are obvious. One way to solve this, of course, would be to make it much harder to get a patent on anything. But the objective of raising patent quality has to be more than just a means to stop bad patents. It has to include a way to make sure that inventors do get patents for truly novel, non-obvious inventions, that such patents are processed relatively quickly and reliably, and that, once granted, they protect subsequent investments in the invention.

**Reduce Uncertainty.** Today, companies and individuals rightly fear that their research will come to naught because someone will assert an as-yet unknown or untested patent against them. Further, when such an assertion of patent infringement is made, the uncertainty about the ability to defend against that assertion often leads to abandonment of the infringing technology.

**Keep Costs Under Control.** The PTO currently spends $1 billion annually. Patent applicants spend several times that amount, and patent litigants spend billions more. Ideally, the PTO’s finances should be decoupled from the amount it raises from fees. Realistically, however, it is unlikely that increased resources will be made available. So we need to look for solutions that go beyond throwing money at the problem.

Some Simple Truths About Reform

**Mistakes Will Always Be With Us.** Patent examiners are human. More important, there is an irreducible element of subjectivity in determining if an invention is truly new. Therefore, we cannot hope to have a system in which no “bad” patents are issued. What is important is to have a system with fewer bad patents, as well as one that functions reasonably well despite the issuance of some bad patents.

It would be very expensive to give all patent applications a sufficiently thorough examination to weed out most bad patents. And while significant increases in spending might be justified, it isn’t likely to happen in the current fiscal environment. But fortunately, it is not necessary.

**Much More Chaff Than Wheat.** Most patents are, and always will be, unimportant. This is not the fault of the patent system, but a feature of the innovation process. For the significance of a new idea usually cannot be known when it is first developed because significance turns on subsequent developments, both technological and economic. And if most patents are doomed to be consigned to the dustbin of technological history, it can’t make sense to spend a lot of resources examining all applications.

Most of what patent examiners do resembles what the referees do in the last minute of a football game where one team is already winning by 30 points. They go through the motions, but their rulings don’t affect the outcome of the contest. The key difference is that in the patent game, usually no one knows whether a particular case is going to matter.

**Rational Ignorance.** In the colorful phrase of Stanford Law School professor Mark Lemley, we can think of the poor quality of patent examination as representing “rational ignorance,” by which he means that society is rationally choosing to remain ignorant about which patents really should be granted by the
PTO. Lemley argues that it is, in fact, reasonably efficient to acknowledge that PTO examination will be of poor quality, and that the cases that really matter will have to be sorted out by the courts. Court cases are expensive. But since only the small fraction of patents that matter will ever get litigated, Lemley argues that litigation is, overall, efficient.

We don’t agree. While the out-of-pocket cost of litigation may be tolerable, the intangible cost of a system with pervasive low-quality patents is much higher than just the cost of paying lawyers to file and defend patent cases. The uncertainty that the current system creates for all parties regarding who may use which technologies is a cost that is very hard to quantify, but is surely significant. Further, this uncertainty undermines everyone’s incentive to invest in new technology. Fortunately, there are changes that could be made that would improve patent quality without requiring dramatic increases in outlays for examinations.

Inventors Respond to How the Patent Office Behaves. The key to more efficient patent examination is to consider how the nature of the examination process affects the behavior of inventors and firms. To put it crudely, if the patent office allows bad patents, people with bad applications are encouraged to show up. If the PTO pretty consistently rejected applications for bad patents, people would understand that applications for bad patents are a waste of money. While some people would still try, the number of applications would surely go down.

A vicious cycle has emerged in which bad examinations increase the application rate, which in turn overwhelms the examiners, reducing examination quality further. If tools could be found to improve patent quality, this feedback process would operate in the other direction, reducing the application rate and freeing resources to further improve quality.

Potential Litigants Respond to How the Courts Behave. When the CAFC issues rulings that increase the chance of the patentee prevailing in an infringement suit, the consequences are not limited to specific cases. Conversations with attorneys involved in patent disputes make clear that the CAFC’s strengthening of the offensive and defensive weapons of patentees has significantly increased their willingness to bring suit. Similarly, the change has significantly decreased the willingness of accused infringers to fight, even when they believe that the patents being used to threaten them are not valid. Constraining the growth
in litigation, as well as the uncertainty created for all innovators by the risk of suit, requires a change in these incentives.

**Get Information to Flow Into the PTO.** Much of the information needed to decide if a given patent application should lead to approval – particularly information about related technologies – is possessed by competitors of the applicant rather than the PTO. And plainly, there are strong incentives for competing firms to share this information. So it would make sense to create opportunities for getting the message through.

But it is important to recognize that private parties’ reactions to the incentives they face can also gum up the works. In particular, any opportunity created for outsiders to provide information that is adverse to their competitors’ patent applications will be exploited opportunistically. This means that any change in procedures that makes it easier for competitors to intervene will, to some extent, increase the cost, uncertainty and delay for valid patent applications.

Ultimately, attention to incentives can mitigate – but not eliminate – the tradeoffs that must be made among the cost of the system, its reliability in screening out bad applications, and the speed and certainty with which good applications are approved. We can’t weed out the trash without killing some good stuff, or greatly improve sorting without expending more resources. But if we pay attention to incentives, we can recalibrate the system to get a better balance between rapid approval of good applications and reliable rejection of bad ones – and do it without dramatically increased resources.

**BUILDING BLOCKS OF REFORM**

There are three key conceptual pieces to our plan for patent policy reform:

First, create incentives for parties that have information about the novelty of inventions to bring that information to the PTO when it is considering a patent grant. Second, provide the possibility for multiple levels of review of patent applications, so that money is not wasted on unimportant patents but sufficient care is taken to avoid mistakes where the stakes are high. Third, replace juries with judges and special masters in ruling on claims of patent invalidity based on the existence of prior art.

The first two pieces are aimed at making the PTO more effective, at a reasonable cost. The third addresses the reality that the best of all possible patent offices would still make
mistakes, so we need a court system that is capable of rectifying those mistakes.

The Quest for Quality at the PTO

Ensuring the quality of the examination process has two key building blocks. First, Congress and the courts must provide the PTO an appropriate standard for issuing a patent – particularly a definition of non-obviousness that separates the wheat from the chaff. Second, the PTO must have the appropriate procedures to implement that standard.

Decisions of the CAFC have made it harder to use identified prior art that renders an application obvious. Nonetheless, we propose beginning with improving the PTO and its processes. It is possible that these reforms will be inadequate – the revamped PTO may feel constrained by CAFC rulings to issue patents despite the existence of important prior art, or a reinvigorated PTO may find its decisions to deny patents repeatedly overturned by the CAFC. If that happens, the issue of reforming court procedures will have to be joined.

Under our proposed revamping of PTO procedures, examination would begin as it does now – with the review of an application by an examiner and without participation by other parties. If, however, the examiner determines that a patent should be allowed, there would be an opportunity for pre-grant opposition – a brief period in which other parties could submit evidence of prior art to make the case that the invention is not novel or is obvious. Pre-grant opponents would not have the opportunity to argue their case, nor would they have access to legal discovery processes. It would simply be an opportunity for parties to inform the examiner.

If, after reviewing the evidence, the examiner decides to issue the patent, there would be a final opportunity for review in the form of a request for re-examination. This request would have to include a stated basis for declaring the patent invalid, and the PTO could decline to grant the re-examination request if no such basis exists. But if re-examination does commence, the procedure would constitute a complete review of the initial decision by an independent examiner, along with opportunity for the party requesting re-examination to argue its case.

The logic of this escalating review process is that most patents would never receive anything other than the most basic examinations. But for those applications that really matter, parties would have an incentive and opportunity to bring information to the PTO, and the PTO would have the opportunity rethink its position.

Some strengthening of the examination received by all patents should complement the institution of effective re-examination and pre-grant opposition. The “rational ignorance” principle implies that it is not efficient to give all patents the kind of scrutiny that they get in litigation or even in a formal re-examination. But the current standard is so low that it is almost surely inefficient. If the PTO approximately doubled the average time spent per patent application and combined this increased initial scrutiny with effective re-examination, it ought to be able to credibly commit to a significant reduction in the issuance of obvious and non-novel patents. This, in turn, would set up a “virtuous cycle,” discouraging applications that are only made today because applicants think they can get away with them.

Unfortunately, there is no way to know how large a reduction in applications would be brought about by credibly tightening the
examination standards. But there is reason to believe it would be significant. Since 1990, when the PTO was converted to its current user-friendly structure, applications have increased from about 150,000 per year to about 350,000 per year. It is likely that a transformation of the examination process making it clear that frivolous applications would be denied would reduce applications to the range of 250,000 per year—roughly one-third less than the current rate. This means that a doubling of examiner effort per application could be purchased for a one-third increase in expenditures.

**Leveling the Judicial Playing Field**

As we have emphasized, there are going to be mistakes in the best of examination processes. So it is important that the court system operates as efficiently as possible to rectify those mistakes, while also permitting owners of valid patents to enforce their rights.

The problem of infringement suits based on dubious patents is greatly aggravated by the legal doctrine that any patent granted is entitled to a legal presumption of validity. Anyone challenging an issued patent must prove by “clear and convincing evidence” that the patent is invalid. This standard is not as high as the “beyond a reasonable doubt” standard for criminal cases, but it is not a neutral standard. By contrast, to win an infringement claim the patentee must only prove infringement by a “preponderance of the evidence.” Given that issues of prior art and obviousness are partly subjective, setting a high standard for proof makes it hard for anyone to be confident that they can invalidate a patent.

Another aspect of the presumption of validity is that the kind of evidence that can be presented to prove invalidity is limited. For example, evidence of how little time was spent by an examiner in consideration of an application is typically not allowed, on the theory that an examiner has to be presumed to have done his job properly.

There is an inherent logic to affording a degree of deference to government decisions made under formal rules. But as we have seen, the process by which a patent is granted is fundamentally different from most other administrative decisions. All interested parties are most definitely not invited to participate in the examination process. There is thus a much weaker logical case for the presumption of validity afforded issued patents.

In any event, the current practice of the PTO is clearly inconsistent with a presumption of validity. Reasonable people do not hold a presumption if everyday observation demonstrates that the presumption is often false. Further, because of the rational ignorance principle, validity can never be a sensible presumption to make about all patents. So it might seem logical to drop the presumption of validity, allow evidence of examiner indifference or incompetence to be presented if it exists, and create a level playing field on which the jury simply decides whether the evidence favors validity or invalidity of a challenged patent.

There is, however, an important reason to maintain some presumption of validity. When a start-up firm goes out to raise money, it is important that patents claimed as the basis for the protecting the firm’s technology are presumed valid. Uncertainty is the enemy of investment, so patents of uncertain validity would be much less effective in providing a base for development of innovations.

Thus, instead of eliminating the presumption of validity, we would change the system to one in which validity is a reasonable presumption. This is why a viable re-examination process is so important. Because of the rational ignorance principle, it would never
be reasonable to assume that the outcome of initial examinations was valid. But if all parties have the opportunity to request re-examination on the basis of factual evidence in their possession, then the presumption becomes reasonable.

If re-examination was never requested, it is indeed reasonable to presume that the patent is valid because the parties most likely to have evidence of its invalidity had an incentive and an opportunity to present that evidence. This is not to say that a patent for which re-examination was never requested is proven valid – only that it is reasonable for there to be a presumption of validity, with all that implies regarding the standard of proof that must be met by a party that ultimately wishes to challenge the patent. And if a patent survives a re-examination, the rational ignorance principle does not apply.

A request for re-examination – which under our proposal is reasonably expensive – combined with a decision by the applicant not to withdraw in the face of such a request, tells the PTO that this is an important application. We should expect the patent office in a re-examination proceeding to devote sufficient resources to “get it right.” Again, this procedure doesn’t prove validity, but it is enough to form a basis for a presumption of validity.

Thus, the existence of a viable re-examination option serves the interests both of parties worried about invalid patents and parties that want the full economic benefit of their valid patents. It helps the former by providing a forum in which appropriate incentives are created for third parties to bring forth relevant facts, and for the PTO to devote the appropriate resources to sifting through those facts. It also helps the holders of truly valid patents; if an effective re-examination procedure exists, the fact that it is not invoked provides a logical basis for presuming that the patent is valid.

**Trial by Jury**

Another complaint is that the right of jury trial stacks the deck in favor of patent holders. The evidence in a patent case can be highly technical and the average juror has little competence to evaluate it. On the surface, the effect of juror incompetence would seem to be neutral. Even if it is neutral, however, having decisions made by people who can’t really understand the evidence does increase uncertainty – and uncertainty explains why many accused infringers settle even when they think they have a pretty good case.

But it probably isn’t neutral; jurors’ inability to grasp technical evidence may interact with the presumption of validity in a way that helps patentees and hurts accused infringers.
When one side has to achieve a reasonably high level of proof, it seems plausible that jurors’ inability to truly understand the evidence being presented acts against the party that must achieve a high standard of proof. If, at the end of the trial, the jurors are simply befuddled by the evidence, the most likely outcome is that they will conclude that neither side has made a convincing case. Thus the “clear and convincing evidence” standard combined with decision making by juries makes it likely that the patentee will win on validity questions.

Regardless of who benefits, one must be skeptical that lay jurors are the best decision makers in patent suits. Of course, the right to a jury of one’s peers is a venerated concept in Anglo-American law. But there isn’t really any sense in which a patent jury is, in fact, a jury of peers. A jury of scientists and engineers – the actual peers of the inventor – probably would be relatively competent. But, of course, that is not what we get. For their part, judges aren’t scientists. But at least they do spend their professional lives evaluating evidence in many different disciplines, and have to develop some ability to sort through it. Further, a judge always has the ability to appoint a “master,” an outside expert in the service of the court to assist.

The CAFC has, in fact, put some limits on the role of juries in patent cases. In particular, it is the job of the judge, not the jury, to interpret the patent’s claims. Thus, when it comes to the issue of infringement, the judge interprets the technical language of the claims before putting the question to the jury. But when it comes to a lack of novelty or obviousness – which the accused infringer must prove with clear and convincing evidence – the jury gets no such help. These decisions are no less technical than those of claim construction. And it would certainly be feasible for the judge to “construe” the novelty and obviousness of the patented invention, just as the judge construes the claims of the patents.

If the PTO were revamped so that a presumption of validity was appropriate, and if the burden of proving invalidity by “clear and convincing evidence” were made feasible by removing the technical determination of novelty and non-obviousness from the jury, patent litigation would be the right last resort when disputes over patent claims could not otherwise be resolved.

There would still be patent suits, and they would still be expensive. Where technology is changing rapidly and claims overlap, there would still be considerable uncertainty about patent rights.

But the pervasive fear that almost any modern product or process is at risk of an infringement claim would be dramatically reduced. And when claims were made based on patents of questionable validity, accused infringers could negotiate from a position from which both parties expect a reasonably competent determination as to novelty and non-obviousness. This should reduce incentives to pay royalties and settle rather than to fight.