CARROTS AND STICKS TO CREATE A BETTER PATENT SYSTEM

By Jay P. Kesant

ABSTRACT

It is widely recognized that the Patent Office grants overly-broad patents because it has deficient knowledge of the relevant prior art, especially in high technology areas with significant nonpatent prior art.

This paper presents five strategies to: (a) increase the quantity and quality of information obtained by the Patent Office from the patentee and her competitors; and (b) create disincentives for patentees to engage in opportunistic behavior by capitalizing on the information asymmetry between patentees and the Patent Office.

The first and second strategies propose changes to our current prior art information disclosure rules based on insights from the economic theory of incomplete contracts. The first strategy proposes that we give patentees the option of presenting an expanded information disclosure statement (IDS) to the Patent Office, disclosing all relevant prior art, including an analysis of how the claims, as filed, relate to this disclosed prior art. If the patentee chooses to exercise this option, the issued patent will be granted a specific presumption of validity with respect to the disclosed prior art. In the alternative, if the patentee does not choose the expanded IDS option outlined above, then the second strategy proposes that we eliminate the presumption of validity for any patent that issues under the current disclosure rules. Ideally, the first and second strategies can be

© 2002 Jay P. Kesant

† Assistant Professor, College of Law and the Institute of Government & Public Affairs, University of Illinois at Urbana-Champaign. I am grateful to Jan Broekman, Rochelle Dreyfuss, Shubha Ghosh, Bronwyn Hall, Justin Hughes, Mark Janis, Mark Lemley, Joe Miller, Craig Nard, Randall Rader, Toshiko Takenaka, Jay Thomas and Gene Quinn for their invaluable comments and criticisms. I would like to thank Greg Aharonian, Steve Kunin, Clarisa Long, and the other participants for their thoughtful critiques of this paper at the Symposium. I would like to thank my colleagues, Tom Ginsburg, Steve Ross, Bruce Smith and the other participants in the University of Illinois Faculty Workshop Series for their helpful comments. I am very grateful to Yoshiaki Aita and Yuji Nakano for sharing their insights regarding the Japanese patent system with me. A special word of thanks to Paul Callister at the University of Illinois Law Library for his dedicated reference support. I would also like to thank Tom Maloney from the University of Illinois, and Leslie Slavich and Jignesh Bhatt from Georgetown University for their very able research assistance. Finally, particular thanks to Mark Lemley for putting together an excellent Symposium that featured numerous valuable contributions.
adopted together. Alternatively, the second strategy can be adopted by itself.

Based on insights from cognitive dissonance and from empirical data regarding oppositions in Germany and Japan, the third strategy proposes that we institute a pre-grant, third party, patent opposition system based on a publication date that is set at 90 days from the issuance of the first Office Action. This proposal creates a mechanism for third parties to participate in the patent examination process prior to its issuance.

Addressing issues in software patents, the fourth strategy proposes that we mandate the use of representational languages in the specification of computer software patents in order to satisfy disclosure requirements and to explicate the metes and bounds of the claimed patent right for software inventions.

The fifth strategy proposes that we create a one-way, pro-defendant fee shifting system if patents are invalidated or revoked in a litigation or opposition proceeding based on certain categories of prior art that are reasonably likely to be discovered by a diligent patentee. This proposal attempts to increase the costs borne by the patentee of engaging in opportunistic enforcement of bad patents.

If these strategies were implemented in concert, we would put in place incentives and mechanisms to create a better informed Patent Office that is more likely to grant patent rights commensurate with innovation and not impoverish the public domain.

**Table of Contents**

I. **Introduction** ................................................................. 765

II. **Specific Proposals** .......................................................... 770

A. The First Strategy ............................................................ 770

1. The Current Prior Art Disclosure System ................................ 770

2. The Patent System as a Two-Stage Bargain ........................... 772

3. A New Prior Art Disclosure Regime ....................................... 773

B. The Second Strategy .......................................................... 775

C. The Third Strategy ............................................................. 776


2. Rethinking Pre-Grant Opposition Systems ............................ 780

D. The Fourth Strategy .......................................................... 783

1. Disclosure Requirements for Software Inventions .................. 784

2. Employing Representational Languages in Software Patent Specifications .................................................. 785

E. The Fifth Strategy ............................................................. 787

1. Review of Fee-Shifting Rationales ...................................... 788

2. One-Way, Pro-Defendant Patent Fee-Shifting ........................ 795

III. **Conclusion** ............................................................... 797
I. INTRODUCTION

Commentators have long complained about the performance of the U.S. Patent & Trademark Office ("Patent Office"). Much of this criticism is directed at the quality of the patents that are granted by the Patent Office. It is widely suggested that the Patent Office issues patents that are either "facially" invalid or broader than the actual innovation disclosed in the patent application. Both problems result from the Patent Office's inability to accurately determine the scope of information that is already in the public domain or is the subject of other patents (i.e., the relevant prior art) when examining patent applications. This is particularly true in areas such as computer software where identifying the relevant prior art is often difficult.

These problems are not necessarily the result of incompetence at the Patent Office. Several commentators have noted that the Patent Office is being asked to perform miracles because it operates under significant budgetary constraints. In the patent community, it is well-known that the amount of time the Patent Office spends examining a patent application, from initial examination to issuance, is approximately the same as the

---

1. When referring to the United States Patent & Trademark Office, I use the term Patent Office instead of the PTO or USPTO to make clear that I am referring to the patent-related activities of the USPTO.


3. The problem of identifying material prior art is particularly difficult in the area of computer software. See, e.g., Julie E. Cohen, Reverse Engineering and the Rise of Electronic Vigilantism: Intellectual Property Implications of 'Lock-Out' Technologies, 68 S. CAL. L. REV. 1091, 1178 (1995) (noting that "in the field of computers and computer programs, much that qualifies as prior art lies outside the areas in which the PTO has traditionally looked—previously issued patents and previous scholarly publications."). For a general discussion of the problem, see ROBERT MERGES ET AL., INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE 1045-47 (2d ed. 2000).

amount of time an attorney may spend searching for relevant prior art in the first week of a patent litigation. As a result, even doubling the amount of time spent by a typical patent examiner would be insignificant when compared to the time devoted to studying prior in litigation, unless the quality of information made available to the patent examiner is improved.

The problems created by a lack of resources are exacerbated by the localized nature of technical knowledge and the social costs of “bad” patents. First, we must consider the nature of the technical and specialized knowledge with which the patent examiner must acquaint herself in every application in order to make a patentability determination. The localization of knowledge pertaining to science and technology is well recognized in a number of disciplines, including information science, knowledge management, and information economics. For example, in his book *Information Anxiety*, Richard Saul Wurman categorizes all types of information as a series of concentric circles radiating out from an individual, with internal and conversational information occupying the innermost circles and general cultural information occupying the peripheral ones. Scientific and technological information occupies one of the inner circles because such information is not widely shared; rather, it is available only to persons working in a specific field or sub-field.

Similarly, researchers in information science and knowledge management have demonstrated that within any technical discipline, technologists form sub-groups referred to as “invisible colleges.” These are loose, but effective, communication networks within which technologists share information with one another. Within each sub-group, the members work out a rich set of customs, habits, mechanisms, and traditions to define the protocol for information collection, including mechanisms for listening and screening out information. Many of these sub-groups are non-intersecting, and hence, knowledge that is most relevant to their technological activities remains local.

5. Lemley, supra note 2, at 1500 (noting that, depending on the art unit, a patent examiner may spend a total of eight to thirty-two hours on a patent application during its two to three year prosecution period).

6. I use this term in the following sense—a patent is “bad” if it should not have been granted by the Patent Office after a reasonable search and review of the relevant prior art.


A third example illustrating the wide recognition of the localized nature of technical knowledge is found in *The Use of Knowledge in Society*, where noted economist Freidrich Hayek recognized that scientific knowledge is very unlikely to be widely dispersed. Instead, it is most likely to be at the disposal of a few particular individuals, the so-called “experts” in that field of knowledge. Further, Hayek persuasively contends that it is unlikely that any single administrative authority could possess all of the information about various facts that are dispersed among various individuals.

From these insights, it is clear that information regarding the relevant prior art for any patent application is most likely to be known only to the patentee and his competitors. Hence, the Patent Office is unlikely to be well informed about the relevant prior art, creating an asymmetry between the patentee’s information and the information possessed by the Patent Office. Consequently, in many cases, especially those areas with significant nonpatent prior art, it is simply not a matter of providing the Patent Office more resources to conduct a more thorough prior art search. Indeed, the patent examiner may not be aware of where to discover the most relevant prior art once she has gone beyond traditional patent databases. Thus, it is not at all surprising that the Patent Office grants invalid or overly broad patents.

As a second preliminary matter, the social costs of improvidently granted patents are numerous. They include the following: (a) opportunistic licensing royalties/fees (including cross-licensing) collected from licensees who may rationally settle for a license instead of resorting to protracted litigation; (b) the disincentive to downstream innovation, i.e., the social cost of abandoned research activities by the patentee’s competitors who may fear infringement; (c) the cost of wasteful designing-around activities by competitors; (d) the cost of rent-seekers, such as venture capital financiers, who may choose to invest in start-up companies based on bad patents, thereby taking away resources from genuine entrepreneurs; (e) the social cost of supra-competitive pricing, in the absence of noninfringing

---

13. Mark Lemley comes to the same conclusion but under a different rationale. See Lemley, *supra* note 2, at 1508-11.
product substitutes, based on bad patents; and (f) the filing and prosecution costs and the subsequent cost of having the courts fix the Patent Office's oversights.

Without significant empirical research,\(^\text{15}\) it is difficult to meaningfully quantify the magnitude of the total social costs of bad patents. Even the more simple task of estimating unnecessary licensing fees is difficult because the value of a license is dependent on factors such as flat payments, reasonable royalties for direct use and subsequent derivative use of the patented technology, and grant-back clauses.\(^\text{16}\) Nevertheless, momentarily setting aside the diminution in public confidence in an administrative system that issues bad patents, improvidently granting extravagant patent rights presents a real concern in a capitalist economy grounded on efficient uses of resources and strong property rights.

The theoretical approach to eliminating the social costs of bad patents is to set the marginal investment in information gathering by the Patent Office to be equal to the marginal reduction in social cost from granting better patents. Many of the proposals presented in this paper attempt to move in that direction.

Recognizing that the social costs of bad patents are significant and that relevant knowledge about the prior art is localized in the patentee and her competitors, we can frame the problems regarding the Patent Office's performance as follows:

(a) How can we increase the quantity and quality of information obtained by the Patent Office from the patentee and her competitors?

(b) How can we create disincentives for patentees to engage in opportunistic behavior by capitalizing on the in-


\(^{16}\) Mark Lemley attempts to tackle this issue and estimates the maximum social cost of licensing holdups to be $443 million, a figure less than the annual patent prosecution costs. See Lemley, supra note 2, at 1515-19.
In this paper, I attempt to briefly answer these two questions. I outline the substance of five strategies that, when implemented in concert, are likely to result in a better-informed Patent Office and a better patent system. I do not present a detailed analysis or defense of any of these strategies, because that is a significant endeavor that I have undertaken elsewhere. Instead, I present these strategies as thoughtful proposals that are worthy of careful consideration.

The first and second strategies propose changes to our current prior art information disclosure rules. The first strategy proposes that we give patentees the option of presenting an expanded information disclosure statement ("IDS") to the Patent Office, disclosing all relevant prior art, including an analysis of how the claims, as filed, relate to this disclosed prior art. If the patentee chooses to exercise this option, the issued patent will be granted a specific presumption of validity with respect to the disclosed prior art. In the alternative, if the patentee does not choose the expanded IDS option outlined above (and either files the current IDS form—PTO Form 1449—or files nothing), the second strategy proposes that we eliminate the presumption of validity for any patent that issues under the current disclosure rules. Ideally, the two strategies are adopted together; as an alternative, the second strategy may be adopted alone. But, as discussed below, it does not make sense to adopt only the first strategy while maintaining the current disclosure system.

The third strategy proposes that we institute a pre-grant, third party patent opposition system based on a publication date that is set at 90 days from the issuance of the first Office Action.

The fourth strategy suggests that we mandate the use of representational languages in the specification of software patents in order to satisfy disclosure requirements and explicate the metes and bounds of the claimed patent right for software.

---

17. For instance, the patentee may seek to enforce an invalid patent and then settle for license fee terms favorable to both the patentee and the licensee, thereby permitting both parties to extract supracompetitive profits from the consuming public. See John R. Thomas, Collusion and Collective Action in the Patent System, 2001 U. ILL. L. REV. 305, 335 (2001).

The fifth strategy proposes that we create a one-way, pro-defendant fee shifting system for situations where patents are invalidated or revoked in a litigation or opposition proceeding based on certain categories of prior art that are reasonably likely to be discovered by a diligent patentee.

The first four strategies attempt to increase the quality and quantity of information that is made available to the Patent Office. The fifth strategy increases the patentee’s cost of engaging in opportunistic behavior by attempting to enforce a bad patent, and reduces the alleged infringer’s costs of invalidating a bad patent.

The remainder of this Article discusses these proposals in greater depth.

II. SPECIFIC PROPOSALS

A. The First Strategy

The first strategy is to change our current prior art information disclosure rules by creating an option for patentees to present an expanded information disclosure statement (“IDS”) to the Patent Office, disclosing all relevant prior art, including an analysis of how the claims are patentable over this disclosed prior art. If the patentee chooses to exercise this option, the issued patent will be granted a specific presumption of validity with respect to the disclosed prior art.

1. The Current Prior Art Disclosure System

At the outset, we must examine the current disclosure requirements governing relevant prior art. The patentee has a duty of candor to disclose prior art that is material to patentability. However, the patentee has no duty to conduct an affirmative prior art search prior to filing a patent application. In other words, the patentee need only disclose any material prior art that is in her possession. This prior art is listed sequentially in an IDS\textsuperscript{19} that is and submitted to the Patent Office, and the patent that issues is virtually bullet-proof with respect to invalidation based on the listed prior art in post-issuance patent litigation.\textsuperscript{20} Note that in the current system the list of prior art references is simply submitted to the Patent Office without stating the relative importance of the prior art references or how the claims, as filed, are patentable over that prior art. This disclosure is of

\textsuperscript{19} PTO Form 1449.
\textsuperscript{20} See John R. Allison & Mark A. Lemley, Empirical Evidence on the Validity of Litigated Patents, 26 AIPLA Q. J. 185, 229-31 (1998) (studying 202 patents over a eight-year period from 1989 to 1996 and finding that the probability of invalidation based on prior art cited in the course of prosecution was very low).
limited assistance to the Patent Office because the patent examiner has the
difficult task of discerning what knowledge is buried in these references in
order to appreciate the import of the disclosed prior art.21

Currently, the doctrine of inequitable conduct22 and the duty of can-
dor23 encourage patentees to produce material prior art to the Patent Of-
fice. Inequitable conduct arises when information material to patentability
is not disclosed to the Patent Office, or when false material information is
disclosed to the Patent Office with an intent to deceive the Patent Office.24
Nevertheless, the level of proof required to establish a claim of inequitable
conduct has been set so high by the Federal Circuit that it has reduced the
incentive to disclose prior art to the Patent Office. The Federal Circuit has
held that even gross negligence in failing to disclose prior art is not in it-
self sufficient to prove an intent to deceive the Patent Office.25 Since 1988,
the Patent Office has indicated that it will not investigate and reject patent
applications based on violations of the duty to disclose material prior art as
it is ill-equipped to make such a determination.26 The Federal Circuit has
frowned on assertions of inequitable conduct as a defense in patent law-
suits and charged that the habit of claiming inequitable conduct has be-
come “an absolute plague.”27 The net result is that there is little incentive
to undertake a complete disclosure of the prior art to the Patent Office. In-
deed, Allison and Lemley report that from 1989 to 1996, nonpatent prior
art was not disclosed to the Patent Office in the vast majority of cases that
they studied.28 Private information regarding the relevant prior art is often
not adequately disclosed to the Patent Office because there is no signifi-
cant incentive for the patentee to do so. As a result, the Patent Office is not

21. I specifically use the term knowledge, as opposed to information, to refer to the
process of internalizing the information found in the disclosed references, thereby turning
the disclosed information into relevant knowledge and permitting assessment of novelty
or nonobviousness of the patent application.

22. See Kingsdown Medical Consultants, Ltd. v. Hollister, Inc., 863 F.2d 867 (Fed.
Cir. 1988); Molins PLC v. Textron, Inc., 48 F.3d 1172 (Fed. Cir. 1995).

23. 37 C.F.R. § 1.56 (1992) (“Each individual associated with the filing and prose-
cution of a patent application has a duty of candor and good faith in dealing with the Of-
fice, which includes a duty to disclose to the Office all information known to that indi-
vidual to be material to patentability as defined in this section.”).

24. See Kingsdown, 863 F.2d at 872.

25. Id. at 873.

26. MARTIN J. ADELMAN ET AL., CASES AND MATERIALS ON PATENT LAW 746

that “the habit of charging inequitable conduct in almost every major patent case has be-
come an absolute plague”).

28. See Allison & Lemley, supra note 20, at 231-34.
equipped to determine the scope of the patentee's nonobvious contribution, particularly in certain technical areas with significant nonpatent prior art.  

2. The Patent System as a Two-Stage Bargain

In order to understand how to improve the current information disclosure system, we must first analyze the current patent system in contractarian terms, relying on insights from the theory of economically incomplete contracts. A general example of a two-stage bargain illustrates the key insights; let us consider a two-stage bargain involving the order and delivery of car doors by a supplier who contracts with an auto manufacturer.

If the auto manufacturer is unsure about the quality of the car doors provided by the supplier, he will reserve, in the first stage, the residual rights to reject the car doors when they are delivered to him in the second delivery stage. If the auto manufacturer preserves these residual rejection rights, he reduces his risk in the event that the car doors are sub-standard upon delivery. In turn, the car door supplier, cognizant of the possibility that the auto manufacturer may reject his car doors, will supply him with generic car doors that can be sold to other auto manufacturers if rejected by the car manufacturer with whom he initially contracted. Therefore, reserving the residual rights to reject the car doors, while somewhat advantageous, will result in the auto manufacturer receiving generic-quality car doors. Stated differently, the division of rights of the contracting parties in the second stage will affect the actions taken by those parties in the first stage and vice versa.

In order to improve the positions of both the auto manufacturer and the car door supplier, the auto manufacturer must clearly state in the contract the standards and specifications for the car doors that he seeks, and, at the same time, relinquish the residual right to reject the car doors if they are delivered in compliance with his specifications. Such a contract allows the car supplier to invest in providing superior car doors, having been contractually assured that they will not be rejected if they comply with the contract. The auto manufacturer is also better off because he now receives a superior product for his car, achieving this result by relinquishing some of his residual rights to reject the delivered car doors. This illustration demonstrates how the contract terms in a two-stage bargain can be modified to improve the positions of both contracting parties. In other words, both par-

29. See Merges, supra note 2, at 589.
ties can engage in a Pareto-improving trade by re-engineering the contract terms in the first and second stages of the contract.\textsuperscript{31}

The patent system can be characterized as a two-stage contractual bargain. In the first stage of this bargain, the prosecution stage, the patentee presents the Patent Office (and hence the public, following the common, but perhaps uncomfortable, assumption that the Patent Office is an agent of the public) an enabling disclosure of her invention and a disclosure of the material prior art. In return, provided the requirements for patentability are met, the patentee receives a contingent set of exclusive rights that are granted by the public. This grant is not unconditional. Rather, the public reserves certain residual rights to invalidate or revoke the issued patent in the second stage, the litigation stage, if the patent is found to be invalid in post-issuance patent litigation. Thus, the public has a valuable \textit{option} that it can exercise in the second stage of the patent bargain.

This option should not be relinquished lightly. In our current patent system, we trade away this option by granting a general statutory presumption of validity to any issued patent,\textsuperscript{32} requiring challengers to produce clear and convincing evidence of invalidity to overcome this presumption.\textsuperscript{33} In return for granting this presumption of validity, the Patent Office receives a list of the relevant prior art. In practice, however, this usually does not enhance the patent examiner's ability to meaningfully examine the disclosure and claims of the patent application at hand. Thus, the current system is a Faustian bargain—the public trades away future residual invalidation rights without receiving in return sufficient current knowledge in return from the better-informed party, the patentee. The public should trade away its residual right to invalidate a patent by granting a presumption of validity only in circumstances where it is economically advantageous to do so.

3. \textit{A New Prior Art Disclosure Regime}

Under this proposal, the public relinquishes its residual invalidation rights by granting a \textit{specific} presumption of validity only when the Patent Office receives an expanded disclosure of the prior art, including a discussion of how the claims, as filed, are patentable over the prior art. This specific presumption of validity is different from the current statutory pre-

\textsuperscript{31} A situation is said to be Pareto efficient or Pareto optimal if there is no change from that situation that can make someone better off without making someone else worse off. A. MITCHELL POLINSKY, \textit{AN INTRODUCTION TO LAW AND ECONOMICS} 7 (2d ed. 1989).

\textsuperscript{32} 35 U.S.C. \textsection 282 (1994).

\textsuperscript{33} EZ Dock, Inc. v. Schafer Sys., Inc., 276 F.3d 1347, 1351 (2002).
sumption of validity, and attaches only to disclosed prior art. Under this specific presumption of validity, a court will not invalidate a patent based on disclosed prior art unless it is convinced that no reasonable examiner would have allowed the patent in light of the disclosed prior art.

Simultaneously, the patentee’s position is improved because the patentee is granted a patent that is less likely to be vulnerable in post-issuance litigation, thereby increasing the patent’s value to potential investors. Indeed, the patentee has every ex ante incentive to maximize prior art disclosure and analysis, knowing that the issued patent will be more immune from attack based on his disclosure of prior art. By creating such an option for patentees, this scheme reduces the information asymmetry between the patentee and the Patent Office, and simultaneously creates an opportunity for the patentee to acquire a property right that is more certain. In short, this is a Pareto-improving trade for both the Patent Office and the patentee.

Furthermore, once such a scheme is in place, patentees who opt into it, and receive patents that are less vulnerable to post-issuance attack, will build patent portfolios that are more attractive to investors. In other words, the market can then value these patent rights appropriately and recognize the significance of the option to obtain a superior patent right, devaluing in turn the patent rights of those who do not wish to choose this disclosure option.

The enhanced prior art disclosure requirements under this scheme will need to be specified and implemented with care in order to maximize the relevant prior art knowledge the patentee reveals to the Patent Office. Once these standards are specified, they may form the basis for additional claims in post-issuance litigation challenging whether these disclosure requirements were actually met in prosecution. However, this is a small price to pay in order to reduce the informational asymmetry between the patentee and the Patent Office. Moreover, administrative agencies, such as the SEC, FDA, and the EPA, routinely require carefully specified information disclosures from regulated firms. Disclosure schemes are how the administrative state’s tool for receiving important information from the better-informed individuals who appear before it.

35. See, e.g., Cynthia A. Williams, The Securities and Exchange Commission and Corporate Social Transparency, 112 HARV. L. REV. 1197, 1210 (1999) (noting that securities laws mandate disclosure to advance a number of purposes such as preventing fraud and promoting market efficiency).
Finally, this first strategy can be implemented as a scheme that patentees can choose to opt into, and the second strategy can be adopted as the default for those patentees who wish to continue with our current disclosure practices. If the first strategy is implemented, but the Patent Office does not appear to benefit from the enhanced disclosures about the relevant prior art and continues to grant overbroad patent rights, then we should consider the second strategy as an exclusive alternative.

B. The Second Strategy

The second strategy is an alternative to the first strategy. If the patentee does not choose the enhanced prior art disclosure option outlined above, any statutory presumption of validity for a patent that issues under the current disclosure rules is eliminated.

If the patentee does not elect to provide an enhanced prior art disclosure, as presented above, the public should retain all residual rights to invalidate the patent in post-issuance litigation, and should not grant any presumption of validity. As has been discussed above, it is economically worse for the public to trade away their post-issuance invalidation rights when they are the less informed party and the disclosure rules do not sufficiently alleviate that information disparity.

Consequently, if the first strategy is not an attractive choice, given the continued concerns that the Patent Office may not be able to appreciate the import of the disclosed prior art even in a world with expanded information disclosure requirements, it makes economic sense to embrace the second strategy exclusively. The general presumption of validity accorded to all issued patents has been interpreted by the Federal Circuit as requiring clear and convincing evidence to invalidate a patent in post-issuance litigation. Implementing the second strategy is equivalent to lowering this burden of proof. Moreover, it is evident from the foregoing that only adopting the first strategy, and rejecting the second strategy—i.e., continuing to preserve the current disclosure system and the general presumption of validity while embracing the first strategy—does not make sense.

Most importantly, either the first or second strategy would present an improvement in economic terms over the present patent regime. Currently, the public relinquishes its broad invalidation rights by granting a presumption of validity that attaches to a list of prior art references with little or no improvement in the Patent Office's ability to examine the patent applica-


37. See *EZ Dock*, 276 F.3d at 1351.
tion under consideration. While eliminating the presumption of validity in these circumstances disadvantages the patentee, as the better-informed party, the patentee can undertake efforts to reduce the likelihood of post-issuance invalidation by performing a careful prior art study.

C. The Third Strategy

The third strategy is to institute a pre-grant, third party, patent opposition system based on a patent publication date that is set at ninety days from the issuance of the first Office Action.

A number of commentators have argued that we should adopt a third-party opposition system, akin to post-grant patent opposition systems in Europe and Japan, in order to bring parties with more knowledge than the Patent Office into the patent examination process, thereby empowering the Patent Office to more accurately discriminate among patent applications. In this section, relying on insights from post-decisional cognitive dissonance theory and on empirical evidence from Germany and Japan, I suggest that any third-party opposition system instituted in the U.S. must be a pre-grant opposition system. In the next sub-section, I briefly review pre-grant and post-grant patent opposition systems and also discuss a recently introduced House Bill that would add an opposition proceeding to the United States patent laws.

1. Review of Patent Opposition Proceedings

As discussed earlier, patent systems where significant numbers of patents are of questionable validity undermine public and investor reliance upon the issued patent as an enforceable property right. Information, especially relevant technical information, tends to remain local, which creates the problem that the patent examiner often will not have sufficient knowledge of the prior art, especially unpublished prior art, and is therefore unable to sufficiently scrutinize a patent application. As a result, a


40. See id.
patent of questionable validity may issue and potentially expensive litigation might follow.

This problem may be ameliorated by adopting an opposition proceeding into the patent examination process. An opposition system would permit third parties to challenge the validity of a patent in an administrative proceeding with the Patent Office. Opposition proceedings may be pregrant or post-grant. A pre-grant opposition system allows challenges to the patent to be made after publication of the patent application but before the patent is granted. A post-grant opposition system allows challenges to the patent to be made after the grant of the patent.

A pre-grant opposition system has the advantage of greater certainty in the validity of granted patents because any challenges to a patent would come before the patent is granted. Additionally, it is less likely that the Patent Office would become entrenched in its position regarding the validity of a patent as a result of fidelity to its examination process. It may be difficult for the Patent Office, which sets out procedures and enforces standards for issuing patents, to turn around and revoke patent claims in a post-grant opposition proceeding immediately after allowing them. Furthermore, if a significant number of oppositions are successful, then the patent examination process could be called into question, a prospect that the Patent Office would not desire. These factors suggest that a more extensive examination process which includes a possibility of pre-grant, inter partes opposition may be more beneficial than a post-grant opposition system.

The pre-grant opposition system requires early disclosure of the patent application. Critics claim that by doing so, the Patent Office would be succumbing to the influence of large corporations, who could begin to use and/or design around a patent at a much earlier point in time. The underlying concern is that, as a result of early disclosure, what was previously a trade secret in the application is now laid bare to the public. If the patent does not issue, the inventor has no means to protect her invention. Propponents of early disclosure systems point to the fact that early disclosure maximizes the economic value of the patent system by promoting disclosure which is likely to lead to further advances in the technology. Propponents also point to the fact that the U.S. patent system remains the only system of significance that does not require early disclosure for all patent applications.

Critics of the pre-grant system also cite the perceived weaknesses of the former South Korean and Japanese pre-grant opposition systems. These particular pre-grant systems were criticized as being subject to abuse by large corporations that hampered patent applications through protracted oppositions, thereby holding off small inventors with competing or blocking inventions. As a result of such concerns and international pressure, Japan and South Korea both replaced their pre-grant systems with post-grant opposition systems.

It is argued that a post-grant opposition system would avoid some of the problems of the pre-grant system while increasing the reliability of patent rights. As a result of the nature of post-grant opposition, early disclosure is not required, thereby avoiding a significant change to the U.S. patent system.

One concern in a post-grant opposition system arises from the fact that the Patent Office that initially issued the patent is expected to adjudicate the validity of that patent in the subsequent opposition proceeding based on primarily the same considerations. The concern is that the examiner is not likely to concede that the outcome of her examination was incorrect and that the patent should not have been granted. However, considering that the goal of the opposition is to promote the reliability of the issued patent right, the reluctance of the Patent Office to invalidate many issued patents would appear to further that goal. Additionally, looking to the Japanese Patent Office ("JPO") Opposition Procedures, European Patent Office ("EPO") Opposition Procedures and the newly introduced House Bill, H.R. 1333, The Patent Improvement Act of 2001, this problem is assuaged by the fact that the opposition panels are to be composed wholly or in the majority by independent members who had no involvement in the initial examination of that patent.

One important distinction between the U.S. Patent Office and the JPO, relating to post-grant oppositions, is the traditional roles of the respective patent offices. The Patent Office has the authority to grant patents, but, after a patent is granted, the Patent Office does not adjudicate the validity of subsequent claims; Section 1338 of the Federal Rules of Civil Procedure gives exclusive jurisdiction of patent cases to the federal district courts. In contrast, the JPO has exclusive jurisdiction over questions of

43. U.S. businesses that felt that there was collusion between Japanese corporations attempting to exclude and delay U.S. patent applications through the pre-grant opposition process.
As a result, the practice of invalidating a previously granted patent is part of the JPO's historical function, whereas granting such a function to the Patent Office would involve a substantial change in its role.

On April 3, 2001, Congressman Berman introduced House Bill 1333, the Patent Improvements Act, which proposes to amend Title 35 of the United States Code by inserting a new chapter after chapter 31. The Patent Improvements Act would add an opposition proceeding to the U.S. patent laws. The Patent Improvements Act requires the Director to establish an Administrative Opposition Panel ("AOP"), comprised of not less than 18 administrative opposition judges ("AOJ"). Additionally, patent examiners may be assigned to assist the AOP, except that a patent examiner may not assist in the review of a patent application that he himself had examined.

Any person would be able to file a request for an opposition to a patent on the basis of 35 U.S.C. §§ 101, 102, 103, or 112. Section 321(b)(1)(A), in conjunction with the amendments to Section 41(a) of Title 35, regarding the fee of $200 for filing an opposition "based on prior art citations or obviousness" and the fee of $5000 for filing an opposition on "any other basis" indicates that the Patent Improvements Act acknowledges the fact that competitors and insiders in the industry are more knowledgeable about the prior art and the obviousness of a patent than the Patent Office. In order for the request to be valid, the request must be made in writing, within nine months of the issuance of the patent, and must be accompanied by the payment of the fee specified in 35 U.S.C. § 41(a).

The Patent Improvements Act would create an expedited administrative procedure for certain challenges to a patent's validity rather than relying on patent litigation in the federal district courts. Section 321(b)(1)(A) requires the Director of the Patent Office to send a copy of the request to the owner of the patent within two months after the request has been made. Next, the patent owner is granted a reasonable period of at least sixty days from the mailing of the request by the Director to "file a statement in reply to the grounds for the request, including any amendment to the patent and new claim or claims, for consideration in the opposition procedure."
processing,” and “the patent owner shall promptly serve a copy of the statement on the third-party requester.52 Within two months of such service, the third-party requester may file a reply to the statement filed by the patent owner.53 The AOJ in the opposition proceeding shall determine patentability of the subject matter of the patent within eighteen months, a copy of which shall be placed in the official file of the patent and promptly mailed to the owner of the patent and the requester.54 Although, this is still a shorter period than in typical litigation, it is a significant period of time.

2. Rethinking Pre-Grant Opposition Systems

The Patent Improvement Act’s proposition of a post-grant opposition system, although a step in the right direction, does not go far enough. Although the conventional wisdom in the patent world favors post-grant opposition systems, I urge the adoption of a pre-grant, third party opposition system with a prompt publication date. I suggest that patent applications be published ninety days after the issuance of the first Office Action. This publication date is carefully chosen to address the concerns associated with an inordinately early publication date, such as prompt publication at the time of filing. In addition, this publication date will permit a patentee to assess her chances of getting an issued patent or preserving the option of keeping her invention a trade secret based on the results of the first Office Action.55

There are several good reasons why any opposition system should be a pre-grant system. Insights from theories of self-perception biases, particularly cognitive dissonance, are compelling in this regard. Post-decisional cognitive dissonance suggests that after the Patent Office has already decided to grant a patent it will be inclined to require a greater quantum of evidence to revoke a patent than is objectively necessary based on the new evidence. As Leon Festinger stated, people experience cognitive dissonance when they simultaneously hold two thoughts that are psychologically inconsistent, i.e., thoughts that feel contradictory or incompatible in some way.56 In these circumstances, people try to reduce cognitive dissonance by systematically discounting the evidence that arouses such feelings. Post-decisional cognitive dissonance follows from the making of a

52. Id. § 321(b)(1)(B).
53. Id.
54. Id. § 321(b)(4).
55. Note that patent applications are examined in secret prior to the eighteen month publication of the patent application. The entire file wrapper for a patent application is made public once the patent is issued.
decision. For example, voters are more likely to believe that the candidate they chose was the best candidate with the best chance of winning after voting than before.\textsuperscript{57}

Applying these insights to patent opposition systems, it is likely that the Patent Office will be more confident that it reached the correct result in granting a patent after deciding to grant a patent, rather than before that decision. This effectively increases the evidentiary burden on third parties and creates a disincentive for them to oppose granted patents. Hence, we should create a system that allows third parties to offer evidence prior to the Patent Office's decision regarding allowance.

The problem of cognitive dissonance may be mitigated by having a different decisionmaker—an independent review board within the Patent Office, for example—for the opposition proceeding, as opposed to the initial patent examination. However, because the second decisionmaker is also drawn from the Patent Office, she is also likely to have an incentive to remain consistent with prior decisions, and to the extent that new prior art may arouse dissonance, she is prone to discount its importance. In other words, she is also likely to demand a greater amount of evidence for revocation of issued patent claims in a post-grant opposition proceeding.

Empirical evidence from Germany and Japan\textsuperscript{58} also supports the contention that pre-grant oppositions are more effective, and indeed, post-grant oppositions are decreasing as a favored option among patent opponents. Data from Germany and Japan for both their pre-grant and post-grant opposition systems suggests that opponents perceive pre-grant oppositions to be more effective in comparison to post-grant systems. In 1981, Germany changed from a pre-grant system to a post-grant opposition system.\textsuperscript{59} In the years 1974 to 1979, there were between 4,600 and 5,250 pre-grant oppositions filed each year. In the same period, nullity proceedings (i.e., court initiated invalidation trials) ranged from 90 to 140 each year.\textsuperscript{60} Since Germany changed to a post-grant opposition system, the number of oppositions has steadily declined from a high of 3,800 post-grant oppositions in 1983 to about 1,650 post-grant oppositions in 1993. In the same time frame, the number of nullity proceedings increased from 93 in 1983

\textsuperscript{57} PLOUS, supra note 56, at 28-29.

\textsuperscript{58} In this paper, I present the empirical data on pre-grant and post-grant oppositions from Japan and Germany very briefly. For a more complete discussion, see Kesan, supra note 18.

\textsuperscript{59} Correspondence from Mr. Ulrich Joos, German Patent Office (1994) (on file with author).

\textsuperscript{60} Id.
to 127 in 1993.\textsuperscript{61} It is important to note that these changes took place in the context of an overall rise in the total number of patents granted each year. The data shows that patents are more likely to be challenged in a pre-grant opposition system compared to a post-grant system.\textsuperscript{62} Once a patent is granted, opponents prefer to challenge a patent through nullity proceedings, i.e., invalidation trials, rather than post-grant oppositions. While there are several plausible reasons for this change, the perception that opponents are more likely to mount a successful challenge to a patent in a pre-grant system seems to have played a role in the decreasing number of challenges in the post-grant system.

It is possible that some of the reduction in the number of post-grant oppositions in Germany may be due to the possibility of filing an opposition at the European Patent Office (EPO) after 1980. However, after an initial surge in EPO oppositions from 1980 to 1987, the percentage of post-grant EPO oppositions has been steadily decreasing from 1987 to 1998;\textsuperscript{63} during the same period, the number of post-grant oppositions in Germany has also sharply declined. Interestingly, the percentage of patent re-examinations in the U.S. has also been decreasing sharply from 1991 to 1998, after remaining relatively stable for many years in the 1980s.\textsuperscript{64}

The data from Japan is similar to the German opposition data and supports the same conclusions. Japan changed from a pre-grant opposition system to a post-grant system in 1994.\textsuperscript{65} Between 1991 and 1995, there were between 5,560 and 8,550 pre-grant oppositions filed each year.\textsuperscript{66} During the same period, there were between 91 and 159 trials for invalidation each year.\textsuperscript{67} Once Japan switched to an ex parte, post-grant opposition system, the number of oppositions dropped from 5,322 in 1997 to 4,558 in 2000.\textsuperscript{68} At the same time, the number of trials for invalidation has in-

\textsuperscript{61} Id.

\textsuperscript{62} Some of the reduction in post-grant oppositions as compared to pre-grant oppositions may be due to the fact that the claims, as filed, were rather broad and were then narrowed in scope during prosecution, thus making it unnecessary to challenge them in the post-grant opposition phase.


\textsuperscript{64} Id.

\textsuperscript{65} Correspondence from Mr. Yuji Nakano, Japanese Patent Office (2002) (on file with the author).

\textsuperscript{66} Id.

\textsuperscript{67} Id.

\textsuperscript{68} Id.
creased from 184 in 1997 to 296 in 2000. Note that the total number of Japanese patents granted each year kept increasing throughout this period. As with Germany, we see that patents are less likely to be challenged in post-grant opposition systems with parties instead choosing to opt for invalidation trials as a means to invalidate patents.

Opponents of pre-grant oppositions are correct in asserting that pre-grant oppositions provide additional opportunities for delay and harassment. For example, large companies with ample resources may choose to stall the issuance of patents by small inventors. However, these concerns can be addressed by procedural devices used to address similar concerns in any litigation or adjudication. The following measures are examples of devices that may limit the possibility of delay and harassment in a pre-grant opposition system: limiting the number of oppositions that may be filed by any third party, forbidding repeated oppositions based merely on cumulative evidence, creating pre-grant oppositions only for anticipatory evidence, etc. Thus, the dangers of delay and harassment should not be the reason for embracing a post-grant opposition regime.

In a pre-grant setting, it is possible that the incentive to challenge a patent may be diminished because the Patent Office may decide to preclude patenting, and as a result, potential challengers may decide not to expend resources until the Patent Office has made a final decision. On the other hand, in a post-grant setting, the Patent Office has already determined that a patent will issue. Faced with the possibility that a patent will issue which may pose a competitive threat, a post-grant challenger may be more inclined to come forward with invalidating prior art. In order to ensure that the incentives to challenge patents in a pre-grant setting are adequately maintained, a list of all the prior art cited by the Patent Office in the first Office Action should be published with the patent application. Consequently, challengers will be able to assess the quality of the prior art they possess compared to the Patent Office’s cited prior art, and thereby, make an informed decision about mounting a pre-grant challenge.

D. The Fourth Strategy

The fourth strategy would mandate the use of representational languages in the specification of computer software patents.
1. Disclosure Requirements for Software Inventions

Patent law sets out general patentability standards that are divorced from any specific area of technology. The process of applying these general patentability standards to an area of technology poses problems that are specific to that technology, including the historical treatment of inventions in that technology and the consistency with which the general standards are policed in that arena.

Computer software is a prominent example. The problems generated by the fact that, until recently, computer software was not patentable subject matter continue to cast a long shadow on the current examination of software patents. Even a decade ago, patent attorneys disguised software innovations as mechanical inventions to work around subject-matter eligibility requirements. As a result, prior art patented software is found in many different art units. Furthermore, patentees used different terminologies (based on their individual organizations) to refer to the same underlying technique. This makes the problem of locating relevant prior art even more difficult. Finally, the English language is a blunt instrument to describe computer software.

Since computer software has been considered patentable subject matter, the need to write around eligible subject matter rules by disguising software inventions as mechanical inventions no longer exists. Therefore, we must avoid the consequences of creating what I refer to as "false translation problems" by allowing patentees to describe their software inventions in the same manner as they would describe it to a fellow software programmer. After all, both the patent examiner and the inventor are most likely educated and skilled in the field of computer software. Thus, there is no reason to disguise computer software inventions.

Moreover, the disclosure requirements, particularly enablement and written description, have been virtually written out of the basic quid pro


quorum between the patentee and the public in the case of software inventions. The Federal Circuit has chosen not to impose specific criteria for software inventions through which the enablement and written description requirements may be satisfied. Indeed, the Federal Circuit has held that high-level, functional descriptions are sufficient to satisfy the enablement and best mode requirements. Thus, for software inventions, general hand-waving amounts to an enabling disclosure. The Federal Circuit has treated the implementation of these functional descriptions in software as a "mere clerical function" for a skilled programmer. Thus, in software patents, there is no requirement that the specific innovation in the software be described to an ordinarily skilled programmer in a manner that explains how the innovation was actually implemented.

2. Employing Representational Languages in Software Patent Specifications

Although software patents can now be sought openly, they are nevertheless inadequately described in patent disclosures. Therefore, the Patent Office should mandate the use of representational languages in the specification of computer software patents. A general-purpose representational language is a language that expresses computer functions in real-world terms. Patent Office examiners who are skilled in computer software will be able to comprehend the inventive step in software inventions more precisely. In due course, we will create a significant repository of patented software prior art upon which the Patent Office can rely.

Mandating the use of representational languages for software inventions also ensures thorough compliance with specification and claim requirements, such as the written description, enablement, and claim definiteness requirements of 35 U.S.C. § 112, ¶¶ 1-2. Note that this approach has long been embraced in chemical and biotechnology inventions; in these technological arenas, use of chemical formulae, listing of nucleotide sequences, and producing deposits and specimens, when necessary, have been routine requirements specified by the Patent Office.

76. The Federal Circuit has explained that "the conversion of a complete thought into a language a machine understands is necessarily a mere clerical function to a skilled programmer." Northern Telecom, Inc. v. Datapoint Corp., 908 F.2d 931, 941-42 (Fed. Cir. 1990) (quoting In re Sherwood, 613 F.2d 809, 817 n.6 (C.C.P.A. 1980)).
A general-purpose representational language is a language that expresses computer function in real world terms. Representational languages convey the structure of a program in a readable English form. They allow an average reader to understand the steps of an algorithm without regard to the specific implementation or platform underlying the function. A representational language may be used to illustrate the elements of a method or algorithm as the interconnection between the elements.

Broadly speaking, representational languages in computer software include object-oriented languages, modeling languages, pseudocode, and knowledge representation. An object-oriented language is a language that organizes software “as a collection of discrete objects that incorporate both data structure and behavior.” This includes details about the language’s structure and rules that enable a programmer to effectively depict the details of her computer program, thereby communicating the essential features of the program at a level that makes reproduction reasonably possible. For example, an object-oriented approach includes organizational constructs such as identity, classification, polymorphism and inheritance to realize effective software representation.

Similarly, a modeling language can employ various types of models such as state diagrams, object diagrams, and data flow diagrams to illustrate a system. Pseudocode is another example of the representation of a computer algorithm in the form of English words and mathematics. While the format may vary depending upon the system and environment, pseudocode allows the programmer to identify basic program concepts and program flow, and it is a standard tool in computer program design.

The Manual of Patent Examining Procedure (“MPEP”) mentions block diagrams and flow charts as methods of describing software process claims, but they are only casually mentioned and no guidance is provided as to what is a “reasonably detailed” flow chart or diagram. As noted earlier, the Federal Circuit has shied away from addressing the sufficiency of the technical disclosure in a flow chart or diagram. A representational

78. Id.
79. Id.
81. Id. at 1.
82. Id. at 1-6.
84. Id.
85. MPEP, supra note 70, at § 2400 et seq.
86. See supra notes 75-76 and accompanying text.
language such as a representational modeling language may be able to provide the structure and detail desired by the Patent Office while furnishing a readable disclosure useful to patent examiners, courts and the general public. Software system designers are already accustomed to modeling and other representational languages in the initial phase of program design. Programmers put their concepts into words and basic steps before implementing them in computer code. Programmers also employ multiple levels of representation before arriving at the final computer language implementation. Hence, it is not overly burdensome for programmers to employ the same language representations in each software patent application. Mandating a technical method of disclosure will increase the transparency of software inventions to patent examiners, the courts, and ultimately, to improvers and competitors.

E. The Fifth Strategy

The final strategy is to create a one-way, pro-defendant, fee-shifting system if patent claims are invalidated or revoked in a litigation or opposition proceeding based on certain categories of prior art that are reasonably likely to be discovered by a diligent patentee.

The patent system is ripe for litigation reform addressing the problems posed by overly-broad patents that are granted by the Patent Office and subsequently enforced by the courts. Fee-shifting is one way to increase the costs of opportunistic patenting and patent enforcement for plaintiffs who choose to strategically assert their invalid patents in litigation. Under such a regime, if a patent were to be entirely or partially invalidated or revoked in a litigation or opposition proceeding, the plaintiff or patentee would have to pay all or a part of the defendant’s fees or the third party opponent’s fees. In addition, the Patent Office could be compensated for the costs associated with examining an invalid patent. Finally, if the plaintiff obtained any monopoly profits based on a patent that was subsequently invalidated in litigation, those profits could be disgorged based on an unjust enrichment theory.

Under this proposal, fee-shifting is implemented when a patent is revoked or invalidated based on certain categories of prior art that are reasonably discoverable by a patentee’s diligent prior art search. Fee-shifting would not apply in situations where a patent is invalidated based on the

87. In this context, the term “defendant” includes a plaintiff in a declaratory judgment action.

88. Mark Lemley has called for a balance in fee-shifting awards to ensure that more accused infringers who prevail also receive fee awards, thereby encouraging more fee-shifting in patent cases. See Lemley, supra note 2, at 1530-31.
sales or other acts of third parties that may not be discoverable when conducting a prior art search.

In exceptional patent cases, 35 U.S.C. § 285 allows a court to grant reasonable attorney fees to the prevailing party. The Federal Circuit has noted that the only deterrent to the improper filing of unwarranted lawsuits on obviously invalid or unenforceable patents is section 285. Nevertheless, the requirement in section 285 of establishing an “exceptional case” remains a formidable barrier that has been difficult to establish. I urge that the underlying fee-shifting contemplated by section 285 be expanded to include situations when patent claim(s) are revoked or invalidated based on certain categories of prior art that may be reasonably discoverable by a diligent patentee conducting a prior art search.

The design of a fee-shifting regime is specifically tailored to an individual application. Because each type of litigation involves many factors, such as the amount of information available to the parties, the chances of winning, and the amount at stake, fee shifting is customized to fit a particular situation and to promote the incentives thought to be desirable. If an appropriate fee-shifting regime were chosen for a particular situation, it could be supported by both fairness considerations, incentive costs, and economic efficiency.

1. Review of Fee-Shifting Rationales

In order to understand some of these rationales for fee-shifting, a brief review of fee-shifting and its rationale is appropriate. Britain and the Commonwealth nations, as well as a number of other countries, operate under a two-way fee-shifting, or “loser-pays,” legal system. The United States, on the other hand, has employed a no-way fee-shifting, or “pay-your-own-way,” system since at least the Eighteenth Century. However, there are many exceptions to no-way fee-shifting in the United States.

State courts in Alaska have long followed a loser-pays rule. Other states, such as Oklahoma and Oregon, have enacted loser-pays rules with mixed results. Florida adopted the loser-pays rule for medical malpractice cases from 1980 until 1985. Some federal laws also make exception to the no-way fee shifting rule, such as the Civil Rights Attorney’s Fees Award Act of 1976, which is a one-way fee-shift that favors civil rights plaintiffs.

There has been much discussion on the merits of the different fee-shifting systems. These debates can be organized generally by fairness considerations and incentive arguments, which include various efficiency considerations. The type of fee shifting employed depends on the rationale, whether it is for fairness, a particular incentive, or a combination of two or more reasons.

Some commentators urge that the loser-pays rule would reduce speculative litigation and limit the tactical leverage gained by a party with a weak case threatening a defendant with costly litigation. This rule may also deter excessive discovery and the filing of unnecessary motions. Some commentators believe that, while the British system would reduce frivolous claims, it would also discourage valid claims. American jurisprudence often imposes indemnity on frivolous claims or defenses. Imposing indemnity on the losing party for nonfrivolous claims may overdeter. Mark Stein suggests a regime under which the attorney for an unsuccessful plaintiff must share the burden of the defendant’s fees. Under this altered contingency fee scheme, attorneys would only take good cases and plaintiffs would not be deterred by the costs as much as with the British rule.

95. Id. at 1164.
96. Id. at 1175-80.
99. Id. at 652.
100. See id. at 667.
101. Olson & Bernstein, supra note 94, at 1161.
102. Id. at 1162.
104. RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 630 (5th ed. 1998).
105. See Stein, supra note 103.
106. See id.
Some believe that the loser-pays rule would streamline proceedings. A claimant or defendant, even if he had only a small chance of losing, would be discouraged from engaging in costly discovery, frivolous motion filing, or, “throwing everything but the kitchen sink,” at the other party. Others, however, believe that there is no reason to believe that either the British rule or the American rule is more efficient overall. It depends on assumptions that a court is not likely to accurately ascertain.

It is difficult to predict whether different fee-shifting rules would increase or decrease the overall rate of litigation. If the plaintiff is more optimistic, he is more likely to bring suit under the British system. If he is more pessimistic, he is less likely to bring suit under the British system. The pessimistic party is discouraged, while the optimistic party is encouraged under the British system. If both parties are pessimistic, settlement will likely occur under either system. However, if the plaintiff is optimistic, he is more likely to litigate because he will get a judgment without incurring costs.

Furthermore, under the British system, the risk averse are less likely to bring suit because, without indemnity, the range of costs to them vary from judgment minus costs, if they win, to the cost of litigation if they lose. With indemnity the range of costs is greater: from judgment, if they win, to the total litigation costs incurred by them and their opponents. Risk aversion may not be very important because corporations are probably not very risk averse. Many personal suits may not be deterred because of contingency contracts because many parties are insured.

An indemnity system will likely have a greater rate of appeals. Plaintiffs have greater incentive to appeal because they have a chance at a much greater benefit—the trial costs plus the judgment. The plaintiff’s risk becomes only the added cost of litigating an appeal. Without indemnity, only the judgment is at stake; thus, it is not worth as much to the party that loses the appeal. Indemnity may cause more court congestion because appellate courts are the “bottleneck of a judicial system.”

107. See Olson & Bernstein, supra note 94, at 1162.
108. See Beckner & Katz, supra note 97, at 207.
109. Id.
111. Id.
112. POSNER, supra note 104, at 628-29.
113. Id. at 630.
114. Id. at 629.
115. Id.
One of the major arguments for different fee-shifting regimes is compliance with the underlying legal rule applicable to a case. In the context of tort actions, Keith Hylton determined that the pro-plaintiff fee-shifting rule results in the greatest compliance with the law because victims have greater access to the courts. Second to one-way pro-plaintiff fee shifting are the American and British rules. Both systems provide nearly the same deterrence, except in situations where a plaintiff has an informational advantage.

The British system promotes more compliance with the law because the American system makes it more costly for a victim to litigate, while the British system provides a greater incentive to litigate for victims that are certain that they will win the suit. A plaintiff knows the defendant violated the law will sue under the British system, because he will not worry as much about the costs. For tort actions, the pro-defendant, one-way rule promotes the least compliance with the law. The pro-plaintiff rule also encourages the least amount of litigation. Even though the pro-plaintiff rule encourages suit, when the issue is in compliance with a standard of conduct, and not strict liability, settlement is also encouraged. If the rate of compliance is high, the probability of success for the plaintiff is low, and thus the plaintiff is more willing to settle. While the pro-plaintiff rule generates the most compliance with the law, it may not be the most efficient rule if it leads to over-compliance. However, Hylton suggests that without values to evaluate the tradeoffs between compliance with the law and litigation cost increases, it is impossible to determine whether pro-plaintiff fee-shifting is economically efficient.

A tradeoff exists between efficiency and equity. A fairness expense exists when innocent defendants are forced to pay the plaintiff’s costs. These insights, when applied to patent lawsuits to address the problem of improvidently granted patents, suggest that a pro-defendant fee-shifting rule is likely to result in greater compliance with the law and reduce the probability of opportunistic enforcement of frivolous patents.

Optimal deterrence is achieved when the penalties are high and the enforcement costs are low because this produces the most compliance at the lowest cost to society. Bruce Hay writes that awarding fees to prevailing

116. Hylton, supra note 110, at 1097.
117. Id.
118. Id. at 1098.
119. Id.
120. Id.
121. Id.
122. Id. at 1100.
plaintiffs is one way to achieve this enforcement system. When a social planner determines the desired level of deterrence, he can choose damages and attorney's fees that would promote this level. Deterrence could remain the same, but costs are reduced.

No deserving theoretical reason exists for believing that the British rule promotes efficient primary behavior. If the analysis is limited to very specific cases, it can sometimes be shown that the British or American rule is more efficient. But, because of the number of factors involved, it is unclear whether the American or British rule is better overall at promoting efficient primary behavior.

As a general proposition, the British rule is less likely to induce settlement than any other fee-shifting rule. The loser-pays rule encourages the most severe strategic demands during negotiation. Settlement is more likely under the American system than under the British system. This is so because the British system reduces the plaintiff's costs, assuming he is optimistic. Because the British rule raises the stakes, it makes litigation more attractive to the plaintiff. Litigation is more attractive to both parties when the defendant's estimate of the plaintiff's victory is less than the plaintiff's estimate of his chances of success.

Weak parties can be made stronger under different fee-shifting regimes. In some areas of the law, as in general tort or contract litigation, there is not necessarily a disparity in the strength of parties, even if there is in individual cases. However, in litigation by a small interest against the government, there is generally a disparity in strength. While the British rule may allow weaker parties to bring suit against stronger parties involving small, meritorious claims, it also discourages large, complex suits that do not have a very high probability of winning. One-way fee-shifting

124. Id. at 509.
125. Beckner & Katz, supra note 97, at 215 (comparing the British and American rules).
127. See id.
128. Hylton, supra note 110 at 1079.
129. Id.
130. Rowe, supra note 98, at 664. See also Olson & Bernstein, supra note 94 (noting that governments that want deregulation can enact one-way fee shifting for individuals against the government; many states have equal access to justice acts that have encouraged small business owners to sue the government).
131. Beckner & Katz, supra note 97 at 216.
helps equalize the strength of the parties when there is generally a large disparity in strength.

Another theory that supports one-way fee-shifting against the government is that it encourages claims that benefit the public. The American rule might deter private litigants from bringing a costly claim that might not benefit the party more than his costs, even though the total benefits to society outweigh the cost to the individual. This type of litigant is sometimes called a private attorney general. Some state courts have required fee-shifting on a private attorney general theory in cases where there was little incentive for governmental enforcement, and the litigation benefited large numbers of people. In contrast, the Supreme Court has rejected this idea absent a statute. Congress has enacted some statutes similar to a private attorney general theory, such as the Civil Rights Attorney’s Fees Awards Act of 1976.

In a close case, it does not seem fair to make the loser pay for all of the legal costs. Although the winner had a slightly better case, and should be compensated for his costs, the inequity of making a reasonable defendant pay all of the legal costs outweighs this consideration. A losing claimant could have been completely reasonable in bringing a claim, but still may have lost. It seems particularly unfair when the loss of a case is due to a court-made change in the law. There is often no way for a losing claimant to know when the law will change. Generally applied, a loser-pays rule is similar to strict liability, and no consideration exists for the losing claimant’s reasonableness. Additionally, close cases may be longer and often more costly to litigate, because they are less likely to be settled. A defendant with a valid claim who pursues litigation will be punished more for a better, but nevertheless losing, claim.

When a party has been wronged, he should be fully compensated. One argument for fee-shifting is that this compensation should include reason-

---

132. Rowe, supra note 98, at 662.
134. Rowe, supra note 98at 663.
137. Rowe, supra note 98, at 655-56.
138. Id. at 656.
139. Id. at 655.
140. Id. at 656.
141. Id. at 670.
able attorney's fees. In American law, the make-whole idea underlies much of the law of remedies. If this make-whole rationale required fee-shifting only if the loser did something wrong, inside or outside of court, then the problem of the close case disappears.

This rationale for making the litigant whole does not apply symmetrically. Few cases would shift fees in favor of the defendant to a claim. Generally, American jurisprudence does not consider initiating litigation to be a legal wrong. If the plaintiff brings a baseless claim, or is guilty of malicious prosecution, the defendant will be awarded his attorney's fees. The same is true for plaintiffs in cases for restitution or declaratory judgment. There is no harm if the only monetary loss to a plaintiff is the cost of going to court. A make-whole rationale would not apply unless the plaintiff is subjected to malicious prosecution.

Many states allow attorney's fees to be included in punitive damages. Damages for attorney's fees can also be justified as punishment for abusing the legal system. However, assessing legal fees as damages for abusive litigation may not be sufficiently punitive. When a claimant abuses the legal system, he not only damages the opposing party; he also burdens the court. The amount of legal fees seems unrelated to the amount necessary to punish a party for abuse of the system.

What amount of attorney's fees should be shifted and what is a reasonable fee? Specific guidelines could be adopted that outline what reasonable attorney's fees should be in a particular circumstance. If full indemnity is too severe as an incentive, some fraction of the reasonable attorney's fees could be awarded. The amount could also depend on the rationale for fee-shifting. For example, punitive fee-shifting has little to do with the cost of reasonable attorney's fees, but if making the defendant whole is the rationale, attorney's fees are the real cost to him. Punitive and deterrent rationales could support double or treble damages. Furthermore, there should be judicial oversight lest the optimistic party with a good chance of winning run up high attorney bills to push the predicted

142. Id. at 657.
143. Id.
144. Id. at 658.
145. Id. at 658-59.
146. Id. at 660.
147. Id. at 661.
148. Id.
149. Id. at 673.
150. Id. at 676-77.
151. Id. at 677.
152. Id.
loser into unfair settlement. Another oversight would be to only partially shift fees so the winner will not run up outrageous fees, because he will incur some expense due to high legal bills.\textsuperscript{153}

Who should pay the costs?\textsuperscript{154} The losing party would usually be liable for the fee, if fee-shifting were employed. If the rationale is to make the injured party whole, it follows that the party liable for the injury should also be liable for the attorney’s fees. However, if the lawyer is blameworthy, some of the fees may be shifted to him.\textsuperscript{155} The parties could bear the costs when they are litigating factual issues that only apply to them; the state could bear the costs when parties are litigating issues of law that the state has been unclear about.\textsuperscript{156}

2. One-Way, Pro-Defendant Patent Fee-Shifting

Applying some of the general insights outlined above to patent litigation, one-way, pro-defendant, fee-shifting in a narrow set of circumstances can be an effective disincentive by increasing the cost to patentees of engaging in certain kinds of opportunistic conduct. Patentees that manage to get overbroad or invalid patent claims granted by the Patent Office by taking advantage of the Patent Office’s lack of knowledge regarding some prior art, may be penalized under such a regime. This may occur when their patent claim(s) are invalidated or revoked in a litigation or opposition proceeding based on prior art that should have been discovered by them through a reasonable prior art search. Their penalty would consist of reimbursing the defendant for his litigation costs, in whole or in part.

This type of fee-shifting increases the risk borne by patentees because their range of possible outcomes is now larger. It also creates incentives for patentees to conduct a thorough prior art search before enforcing their patent in court, and at the outset when filing for patent, in order to ensure that they are issued claims commensurate with their innovation. In addition, pro-defendant fee-shifting creates disincentives for patentees seeking to enforce frivolous patents through the courts. For defendants, this type of fee-shifting creates incentives not to settle prematurely if they believe their invalidation case is strong, because their litigation costs may be borne by the patentee. In the case of oppositions, one-way, pro-opponent, fee-shifting encourages parties to successfully oppose published patents.

\textsuperscript{153} Olson & Bernstein, \textit{supra} note 94, at 1162-63.
\textsuperscript{154} Rowe, \textit{supra} note 98, at 672.
\textsuperscript{155} \textit{Id.}
\textsuperscript{156} \textit{Id.} at 672-73.
It is not altogether clear that the pro-defendant, fee-shifting rule would reduce frivolous patents; perhaps it would only reduce their enforcement. But such a rule, even if limited to invalidations based on discoverable prior art, would marginally affect innocent parties by lowering enforcement of valid patents. The more risk-averse the patentee, the more he will be affected by a pro-defendant (i.e., pro-alleged infringer) fee-shifting rule. The difference in outcomes for the patentee is greater with this fee-shifting rule: from patentee's judgment minus his costs to his costs plus the defendant's costs, \{[J-C_p] to [C_p+C_{d}^p]\}. Without fee-shifting, the patentee's outcome ranges from \{[J-C_p] to C_p\}. Similarly, with this fee-shifting rule, the difference in outcome for the defendant varies from \{[J+C_{d}^d] to 0\} and is more favorable compared to the no fee-shifting scenario in which the defendant's outcome ranges from \{[J+C_{d}^d] to C_{d}^d\}. Therefore, unless carefully tailored to circumstances that the patentee believes he can guard against such as invalidation based on discoverable prior art, the possible loss to the risk-averse patentee is greater with the pro-defendant fee-shifting rule. In other words, such a rule may also affect innocent patentees by lowering the enforcement of valid patents unless narrowly tailored to the circumstances noted above. Higher costs will also affect poor or weaker parties more than strong parties. A company with more money will be more willing to take on the extra costs of getting a patent enforced under a one-way, pro-defendant fee-shifting rule.

It may be argued that this one-way, fee-shifting strategy should also be employed when an accused infringer prevails in a patent litigation by establishing that there was no patent infringement by her accused product. The underlying rationale for such fee-shifting is to encourage defendants to continue to litigate and not prematurely settle in order to define the proper scope of protection accorded by the patent. In the same way that an accused infringer who invalidates a patent claim performs a public service by removing an improvidently granted patent, an accused infringer who establishes noninfringement also performs a public service by delineating the proper scope of patent protection. However, I would suggest that we not employ fee-shifting for establishing noninfringement because it is difficult to distinguish between genuine, but perhaps unsuccessful, enforcement of a patent right and the harassment of an accused infringer. I prefer to restrict fee-shifting to those cases where there is clear fault, i.e., the plaintiff is attempting to enforce a patent that he would have realized is invalid had he conducted a diligent prior art search. Fee-shifting in these circumstances creates an incentive for the patentee to conduct a diligent prior art search prior to enforcing her patent rights. Not doing so may result in additional costs imposed by the fee-shifting regime. In addition, in
cases of clear noninfringement, there are other procedural mechanisms, such as moving for summary judgment for noninfringement, that are available to limit litigation expenses.

Pro-defendant fee-shifting does seem to have a good chance of decreasing infringement actions overall, lowering court congestion and speeding up proceedings by encouraging fewer frivolous motions on the part of the patentee. But it may encourage frivolous motions on the part of the alleged infringer. In addition, on the margins, with some risk-averse plaintiffs, it may encourage settlement. However, it is also likely to lower the defendants’ incentive to settle based on the prospect of fee shifting in their favor, if the defendants believe their invalidation case is strong.

III. CONCLUSION

It is widely recognized that the Patent Office grants overly broad patents since it has deficient knowledge of the relevant prior art, especially in high technology areas with significant nonpatent prior art. In this paper, I propose a number of strategies to both create incentives for patentees and their competitors to reveal prior art information to the Patent Office, thereby increasing the quantity and quality of prior art information available to the Patent Office, and to create disincentives to engage in opportunistic patenting and enforcement of frivolous patents. These strategies take the form of a threat of punishment and an offer of reward and they include: expanding the prior art disclosure rules and linking them to a specific presumption of validity granted to disclosed prior art and eliminating the current, general presumption of validity for all issued patents; creating a pre-grant opposition system based on a patent publication date of 90 days from the first Office Action; requiring the use of representational languages in software patent specifications; and creating a one-way, pro-defendant fee-shifting rule in patent litigations or oppositions in which a patent is invalidated, in whole or in part, based on reasonably discoverable prior art. If these strategies were implemented in concert, we would put in place incentives and mechanisms to create a better-informed Patent Office that is more likely to grant patent rights commensurate with innovation and not impoverish the public domain.