On the Temporal Reach of Patent Scope: A Multidimensional Model of Invention-Space

Kevin Emerson Collins*

Later inventors often improve upon products disclosed by earlier patentees, creating new products that were not fully possessed or enabled by those earlier patentees. Improvement raises the question of the *temporal reach* of the earlier inventor's patent, i.e., the reach that an earlier patent should have into future improvements that are not yet invented at the time the patent is filed. Temporal reach raises two questions. The first is the "when" question: Under what conditions should courts construe a patentee's rights to encompass an improvement? The second question is the "how" question: By what doctrinal mechanism should that reach be created? The "when" and "how" questions are usually addressed in distinct literatures, the former employing economic reasoning largely divorced from the reality of contemporary patent doctrine and the latter couched in doctrinal reasoning largely divorced from the economics of improvement. This Article makes an argument that bridges the gap.

There is an unacknowledged blind spot in the contemporary scholarship and doctrine that addresses the temporal reach of patent scope into improvement. The stories that convey the conventional wisdom on a patent's temporal reach all involve *classic improvements*, but this Article demonstrates that another, distinct category—*orthogonal improvements*—is also common as an empirical matter. Orthogonal improvements have been overlooked because improvements have been differentiated only on a linear scale of importance or social value. A *multidimensional model of invention-space* is needed in which the two increments of progress achieved by an earlier inventor/patentee and a later improver do not always relate to each other in the same way. Sometimes later advances refine earlier advances (yielding classic improvements), but sometimes later advances branch off in a different direction and are simply compounded with the earlier advance in the improvement-as-thing (yielding orthogonal improvements).

With respect to the "when" question, orthogonal improvements demonstrate that conventional wisdom is incomplete. The optimal temporal reach for classic improvements is legitimately contested, but orthogonal improvements are easy cases. Earlier patents should, and usually do, have sufficient temporal reach to encompass them. The key is to recognize the successive technological advances that lead to orthogonal improvements are economic complements, just like earlier- and later-developed things can be economic complements. This conclusion, in turn, has significant implications for the "how" question. Here, the problem is not that conventional wisdom is incomplete but rather that it is built on faulty premises. The very nature of peripheral claims needs to be reconceptualized. The literal scope of a peripheral claim—and not only the protection provided by the DOE—does and should routinely expand over time to encompass after-arising improvements.

---

* Kevin Emerson Collins, Professor of Law, Washington University School of Law.

On the Temporal Reach of Patent Scope: A Multidimensional Model of Invention-Space

Kevin Emerson Collins*

INTRODUCTION

Technological progress is a cumulative endeavor, but precisely how patent protection should be structured to most effectively promote cumulative progress is the subject of a robust debate.[1] In particular, patent protection for a product that improves over time poses conceptual difficulties. Later inventors often improve upon the product disclosed by an earlier inventor, creating a new product that was not fully possessed or enabled by the earlier patentee. In this situation, the later actor is clearly entitled to patent protection for the improvement,[2] but there are many divergent views on the optimal *temporal reach* of the earlier inventor's patent, i.e., the reach that the earlier

---

* Kevin Emerson Collins, Professor of Law, Washington University School of Law. I thank Mark Janis and Scott Baker for their (unsurprisingly) valuable comments in preliminary discussions about this paper. The title is an intentional reference to Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990).

[1] [string cite]

[2] Provided the improvement satisfies the nonobviousness requirement, of course. 35 U.S.C. § 103 (2006). The strength of the nonobviousness requirement is an important policy lever in fine tuning patent incentives for improvement, but it lies beyond the scope of this Article.

patent should have into future improvements that are not yet invented at the time the earlier patent is filed.

The optimal temporal reach of a patent raises two distinct questions. The first is the "when" question: When, or under what conditions, should courts construe a patentee's rights to encompass an improvement? Here, economic reasoning about patent incentives and transaction costs usually predominates. Common points of contention include the desired magnitude of patent incentives, the appropriate distribution of incentives among different generations of inventors, the friction in the market for patent licenses, and the relative virtues of competition and coordination in the search for technological progress.[3] The second question is the "how" question: Presuming that some patents reach into later-developed improvements some of the time, how, or by means what doctrinal mechanism, should that reach be created? The crux of the problem here is that the earlier patents must have *depth* in addition to *breadth* if after-arising technologies such as improvements are to infringe. Claim scope must grow over time as post-filing progress ensues to encompass a set of things that is larger than set of things disclosed in a patent specification.[4] The issue of claim depth leads to a debate that is predominantly about legal form.[5] The contemporary patent regime is a *peripheral* claiming regime in which the outer boundary of a patent claim is specified as of the date on which the patent is filed. To the extent that the scope of a peripheral claim is fixed on the date of filing, the notion that the claim can encompass an improvement that is only first invented on a later date is usually seen as problematic at best and a logical impossibility at worst.[6] To head this problem off at the pass, courts and commentators often argue that the doctrine of equivalents (DOE) should be a patentee's sole doctrinal recourse to obtain any claim depth. The DOE extends a patentee's rights beyond the literal scope of her peripheral claim. It makes no pretense of fixing its reach *ex ante*, so it is free to encompass later-developed technologies such as improvements.

---

[3] *See infra* Part II.B (discussing the conventional wisdom on the temporal reach of patent scope into improvement).

[4] *See infra* Part I.B (distinguishing breadth and depth as two dimensions of claim scope). As the term is used in this Article, "improvements" are *not* situations in which an earlier patentee claims thing A, a later inventor invents thing B, and the consumer desires the thing A+B. Claim depth is not required for infringement in this components scenario. Rather, improvements occur only when an earlier patentee discloses and claims thing A and a later inventor invents thing A'—a thing that is distinct from A and does not simply include A as a component. *See infra* Parts I.B & C.

[5] In particular, the debate usually involves a variant of the well known rules-versus-standards debate. Of course, the rules-versus-standards debate and its application to intellectual property claims can be framed in economic terms, too. Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992) (addressing rules and standards); Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. Chi. L. Rev. 719, 727 (2009) (addressing central and peripheral claims in intellectual property).

[6] *See infra* Part III.A.

The "when" and "how" questions of the temporal reach of a patent are usually addressed in distinct literatures, the former employing economic reasoning largely divorced from the reality of contemporary patent doctrine and the latter couched in doctrinal reasoning largely divorced from the economics of improvement. This Article, however, makes an argument that bridges the gap. The argument is grounded on the observation that there is an unacknowledged blind spot in the contemporary scholarship and doctrine that addresses the temporal reach of patent scope into improvements. An entire category of empirically common improvement is right under our noses, but we have overlooked it. The stories that convey the conventional wisdom on a patent's temporal reach all involve *classic improvements*, but this Article demonstrates that another, distinct category—defined below as *orthogonal improvements*—is also common as an empirical matter. Putting the existence of these overlooked improvements openly on the table has significant implications for both the "when" and "how" questions of temporal reach. With respect to the "when" question, orthogonal improvements demonstrate that conventional wisdom is incomplete. The optimal temporal reach for classic improvements is legitimately contested, but overlooked improvements are easy cases, and earlier patents should, and usually do, have sufficient reach to encompass them. This part-descriptive, part-normative conclusion, in turn, has significant implications for the "how" question. Here, however, the problem is not that conventional wisdom is incomplete but rather that it is built on faulty premises. The very nature of peripheral claims needs to be reconceptualized to accommodate the fact that the literal scope of a peripheral claim—and not only the protection provided by the DOE—has depth and routinely expands over time to encompass after-arising improvements.

Contemporary theory and doctrine about the temporal reach of patent scope into improvement are premised classic improvements. A classic improvement arises when successive technological advances build on each other in the sense that the improver must stand on the shoulders of an earlier inventor in order to be in a position to make her contribution to technological progress.[7] The improver's technological advance refines the patentee's technological advance in a way that makes the improvement continue to be an embodiment of the earlier patentee's ideas at a high level of generality but not at a lower level of generality. Classic improvements are the stuff of patent lore, and they monopolize the academy's attention.[8] Early in the development of the automobile, George Selden invented and patented a combination of a "carriage body," a power train, and a "liquid hydrocarbon gas engine of the compression type." Many later inventors, including Henry Ford, refined Selden's invention. In their improvements, they retained Selden's general idea of combining the specified elements into an automobile. In this sense, they stood on Selden's shoulders. However, they supplanted Selden's earlier ideas about the specifics of light, gasoline-powered engines and the means of connecting the

[7] The standing-on-shoulders metaphor is a staple in patent-related discussions of cumulative innovation. Suzanne Scotchmer, *Standing on the Shoulders of Giants*, 5 J. ECON. PERSPECTIVES 29, 30 (1991).

[8] The improvements on the Wright and Selden technologies discussed below are both featured as examples of "cumulative technologies" in one of the seminal articles on improvement. Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 888−91 (1990).

elements of the combination with more effective ideas of their own.[9] Glen Curtiss' developments in the aviation industry subsequent to the Wright Brothers' pioneering invention are classic improvements, too, because they involve the simultaneous appropriation and revision of the Wright Brothers' work. Generally framed, the Wright Brothers developed and broadly patented a means of stabilizing an aircraft by *inter alia* raising and lowering different portions of the wing surface—the "normally flat aeroplane" in the language of patent—at the same time. More specifically, they accomplished this with a flexible wing frame that allowed the entire surface of the wing to be warped. Curtiss stood on the Wright Brothers' shoulders in the sense that he, too, employed a mechanism for stabilizing an aircraft that involved raising and lowering different parts of the wing at the same time. Armed with this idea, however, Curtiss produced a more specific idea that supplanted the Wright Brothers' earlier-developed specific idea. He abandoned wing-warping and invented "ailerons"—discrete flaps on a wing that could move independent of the wing.[10]

Not all improvements, however, are classic improvements. Not all improvements involve successive advances in which the later advance refines the earlier one in the sense that the later advance supplants the specifics of the earlier advance. Sometimes, sequential inventors stand side by side, not foot on shoulder, and they produce improvements in which the successive advances compound with each other and do not supplant each other. For example, assume that the Selden patent is followed by an advance in the art of welding the frame of the carriage body, leading to the sale of an improved automobile that otherwise is the exact embodiment that Selden disclosed in his patent. In the context of the Wright Brothers' patent, imagine that a later inventor makes a far improved canvas to stretch over the surface of the frame of a warping wing. In these improvement scenarios, the contribution of the later inventor to technological progress is not so much a refinement of the earlier inventor's ideas but rather a discrete idea built on an independent technological foundation. Furthermore, the embodiment of the later-developed advance in the improvement does not supplant the technological advance attributable to the earlier patentee in any meaningful way. The improved thing does not embody the earlier patentee's technological advance any less than the embodiments actually disclosed by the earlier patentee did.

This second category improvements is clearly after-arising technology, just like classic improvements are after-arising technology. The improvements are not disclosed by the earlier patent specification. The earlier patents must therefore have depth if the improvements are to infringe: the terms "carriage body" and "normally flat aeroplane" must refer to more things later in the patents' terms than they did at the time the patents

[9] Whether the Selden patent had sufficient temporal reach to encompass the classic improvements of Ford and others was an issue that was litigated for many years and that was decided differently in different courts. Electric Vehicle Co. V. C.A. Duerr & Co., 172 F. 923, 926 (C.C.S.D.N.Y. 1909); Columbia Motor Car Co. v. C.A. Duerr & Co., 184 F. 893 (2d Cir. 1911).

[10] The patent fight between the Wright Brothers and Curtiss is legendary. *See, e.g.*, Wright Co. v. Herring-Curtiss Col., 204 F. 597, 614 (W.D.N.Y. 1913), *aff'd* 211 F. 6545 (2d Cir. 1914).

were filed. Yet, these improvements are neither represented in patent lore nor discussed in patent theory. They have been overlooked. Their absence from patent discourse can perhaps be excused as a practical matter because, unlike classic improvements, these overlooked improvements present easy, uncontroversial cases in which sufficient temporal reach for infringement is the desired outcome. Judges routinely grant, and commentators routinely presume, sufficient temporal reach for infringement.[11] The relevant rule in the hypothetical improvements based on the Selden and Wright Brothers inventions is that a patent on a mechanical invention can encompass devices made out of any after-arising material.[12] This rule reflects the common-sense position that a patent on a doorknob encompasses after-arising, improved doorknobs made out of newly invented materials because a "doorknob is a doorknob"—i.e., it is still the inventor's invention—regardless of the material out of which it is made.[13] For the same reason, apparatus claims to software inventions routinely encompass improvements in which the software runs on after-arising hardware.[14] Because judges and commentators have systematically overlooked these improvements, however, the interesting theoretical questions that lie at the heart of the issue of the temporal reach of patent scope into improvement improvement—In what way are overlooked improvements different from classic improvements? Why is this difference important as a normative matter? How do peripheral claims achieve the depth required to encompass them?—have never been asked. This Article takes up these questions, introducing a new conceptual framework for classifying improvements in order to do so.

The distinction between the two types of improvements derives most fundamentally from the nature of things and the role that things play in the patent regime.[15] Although patent rights are intended to promote the generation of new

---

[11] However, they often do not acknowledge the temporal reach that they grant or presume. *See infra* Part III.C.3 (noting that orthogonal improvements literally infringe even means-plus-function claims that are governed by an express doctrinal prohibition on reaching into after-arising technology).

[12] *See* Robin C. Feldman, *Rethinking Rights in Biospace*, 79 S. CAL. L. REV. 1, 28 (2005) (arguing that the "footprint" of claims to "bioscience" inventions should not encompass AAT) (noting the "one embodiment" rule of enablement that applies whenever claims in the mechanical arts are allegedly overbroad because they encompass devices made from many different materials); Michael J. Meurer & Craig Allen Nard, *Invention, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 GEO. L.J. 1947, 1976–77 (2005) ("An inventor, familiar with [a] trend [toward lighter and stronger rackets], should describe the material used to make his racket in general terms, and then the patent claim will literally cover a racket of the same shape and dimension even if it is made from a substance that was not known at the time of the patent application."). A claim in the mechanical arts cannot read on any material if materiality is a limitation of the claim.

[13] Feldman, *supra* note 12, at 3.

[14] *See infra* text accompanying notes [Part II.D.1].

[15] Things are under-theorized conceptual entities in patent scholarship. *See* Michael J. Madison, *Law as Design: Objects, Concepts, and Digital Things*, 56 CASE W. RES. L. REV. 381 (2005) (mulling on the importance of things in intellectual property in general); Kevin Emerson Collins, *The Reach of Literal Claim Scope into After-Arising Technology: On Thing Construction and the Meaning of Meaning*, 41

technological ideas, they do not treat the newly discovered technological ideas *per se* as the resource that is governed by property rights. Thinking about a pill with a newly conceived shape that makes the pill easier to swallow—or talking about the shape or copying diagrams that represent it—is not infringement. Rather, to achieve the goal of promoting the production of new technological ideas, patents grant inventors rights to exclude others from making, using, selling, offering to sell, or importing a set of things that embody those new ideas and that are described by claims.[16] Things are complex entities, and what it means for a thing to embody an idea is yet more complex. What is clear, however, is there is an important difference between the way in which classic and overlooked improvements embody the earlier and later advances that are wound up in stories of earlier patents and later improvements. The conventional wisdom on improvement assumes that improvement can be measured on a linear scale of importance or social value. How valuable is the later advance that leads to the improvement in relation to the earlier advance that underlies the patent?[17] This linear metric fails to register a normatively important fact that distinguishes two categories of improvement. The two increments of progress in things achieved by an earlier inventor/patentee and a later improver cannot always be charted on a single dimension because the two successive advances do not always relate to each other in the same way. Rather, *invention-space*—that is, the metaphorical space in which the technological progress of things can be mapped—*is multidimensional*. Classic improvements arise when the earlier and later advances involve progress along branching paths that implicate the same dimension of invention-space, with the later advances refining—i.e., in part building on and in part supplanting—earlier ones. In contrast, the overlooked improvements involve successive advances on distinct and independent dimensions of invention-space. The later advances branch off in a different direction from the earlier advances, and the two advances are simply compounded together in the improvement-as-thing. It is for this reason that this Article refers to the overlooked improvements as *orthogonal* improvements.

The notion of a multidimensional invention-space points to an otherwise obscured explanation of why patent scope should routinely have sufficient temporal reach to encompass orthogonal improvements. The key is to identify the properties embodying advances along independent dimensions of invention-space as complementary goods and an orthogonal improvement as a single, indivisible thing that naturally bundles the

---

CONN. L. REV. 493, 514–36 (2008) [*Things and Meanings*] (defining the process of "thing construction" and demonstrating its import in patent law).

[16] Patent law protects methods as well as things. An analysis of earlier-claimed and later-improved methods that roughly parallels the analysis of earlier-claimed and later-improved things is feasible, but, for brevity, this Article addresses only things and leaves methods for a subsequent project.

[17] *See* Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 1070 (1997) (arguing that rights to improvements should turn on "the *relative* importance of the original invention and the improvement as the critical factor"); Merges & Nelson, *supra* note [], at 909–11 (same).

complements together.[18] When successive inventors produce complementary things, such as hammers and nails, the economic ramifications are familiar and intuitive. If an earlier inventor patents a hammer and a later inventor patents a new nail that does not bend as easily, the two patents block the hammer-and-nail bundle. The later invention increases the social value of the earlier invention. Whether the earlier inventor's private return increases is a complicated issue, but it should under no circumstances decrease.[19] When successive inventors produce properties of things that are complementary, the economic ramifications should be identical even if the two properties can be embodied simultaneously in the self-same thing as an improvement. There is no reason to discriminate between complementary properties of a single thing and complementary distinct things; patent rights in the improvement-as-thing that bundles the two advances should be the same as the patent rights in the bundle of the two distinct things. The only way to do this is to grant the earlier issued patent sufficient temporal reach to encompass the orthogonal improvement, i.e., the bundle of the two properties that embody the successive advances.[20] Failing to allow earlier patents to encompass orthogonal improvements would be economically akin to saying that it is not an infringement of an earlier patent on a hammer to make, use, or sell hammers that are employed only to pound on after-arising nails.

Orthogonal improvements also have serious implications for the conventional wisdom on the "how" question of patent depth because they already routinely fall within *literal* claim scope in the contemporary patent regime. Claims to mechanical arrangements of parts routinely literally encompass devices made out of after-arising materials; apparatus software claims routinely literally encompass after-arising hardware that is programmed with the software.[21] This routine literal infringement of after-arising technology has been overlooked, just as orthogonal improvements themselves have been overlooked. Once openly acknowledged, it undermines the premise that stable peripheral claims must, by definition, fix the set of claimed things on the date on which a claim is filed. Rather, it reveals peripheral claims to be claims that fix the outer boundary of a claim on its date of filing without finalizing the membership roll of the set of things that

[18] To date, the concept of economic complements has only been used in a manner that is extrinsic to patent law proper, namely to assess the pro- or anti-competitive nature of patent licenses. *See infra* text accompanying notes [Part III.D.3.a]. Here, it is employed as an intrinsic part of the patent doctrine that determines claim scope.

[19] *See infra* text accompanying notes [Part III.D.3.a].

[20] In contrast, a classic improvement is a part-complement, part-substitute hybrid. *See infra* text accompanying notes [Part II.D.3.b].

[21] After-arising orthogonal improvements even appear to literally infringe means-plus-function claims, despite the express prohibition in the doctrine of means-plus-function claiming on after-arising technologies as statutory equivalents. *See infra* Part III.C.3. The fact that orthogonal improvements can infringe claims under a doctrine that facially prohibits infringement by after-arising technology demonstrates exactly how overlooked orthogonal improvements really are. They often are not even recognized as after-arising technology at all.

falls within the outer boundary on that date. The key to this conceptual shift lies in understanding the nature of how descriptive terms acquire meaning and what it means for meaning to be fixed.[22] The philosophy of language has shown that there can be play between the meaning of language and the set of things that the language describes that allows meaning to be fixed *ex ante* and the set of described things to expand over time. It demonstrates that there is no logical contradiction or irresolvable paradox in the notion of a peripheral claim that is both fixed and capable of expansion. This insight from the philosophy of language reconceptualizes the notion of a peripheral claim as a claim that fixes an outer boundary without fixing the content of what lies within that boundary. This revised concept of a peripheral claim, in turn, has game-changing implications for nearly all patent doctrines that affect the temporal reach of patent scope into improvement—including claim construction, the disclosure doctrines of enablement and written description, the rules of means-plus-function claiming, and the doctrine of equivalents. It also has implications for the ongoing academic debate over the merits of peripheral and central claiming.

This Article proceeds in three parts. Part I lays some groundwork. It distinguishes two dimensions of claim scope that affect the temporal reach of a patent: *breadth* and *depth*. It defines the concept of an "improvement" as it is employed in this Article to be a later-developed thing that can infringe an earlier-filed claim only if the claim has depth. Part II distinguishes classic and orthogonal improvements, and it draws out the importance of the distinction to the "when" question of the temporal reach of claim scope into improvements. It ends with a coda that revisits the seminal article *On The Complex Economics of Patent Scope* by Robert Merges and Richard Nelson and argues that the classic/orthogonal distinction articulated here offers a new perspective on the cumulative/distinct distinction in Merges and Nelson. Turning to the "how" question, Part III redefines the nature of a peripheral claim in a manner that explains how literal claim scope can achieve the depth required to routinely encompass after-arising technology such as orthogonal improvements. It then explores the implications of this revised theory of the peripheral claim, together with the normative conclusion of Part II that classic and orthogonal improvements require differential treatment, on patent doctrine and theory.

I. CUMULATIVE INNOVATION, IMPROVEMENT AND PATENT DEPTH

This Part lays the conceptual groundwork for the arguments made in Parts II and III. Part I.A explains the basic principle of patent protection under which incentives to generate new ideas are created by granting inventors rights to exclude others from sets of things that embody those ideas. Part I.B distinguishes two dimensions of claim scope that can affect the temporal reach of patent scope into improvements: claim breadth and depth. Part I.C defines an improvement and demonstrates that a patent must have depth to encompass an improvement. One purpose of the arguments presented in this Part is to

[22] This argument is explored at greater length in Collins, *Things and Meanings*, *supra* note [], at 536–53 (distinguishing ideational and denotational meaning and demonstrating the importance of this distinction in claim construction).

narrow the topic by flagging issues, bracketing them, and setting them aside. There are many important issues involved in tailoring patent protection to accommodate cumulative innovation, but many of them are beyond the scope of this Article's focus on the temporal reach of patent scope into improvement.

A. Patents, Ideas, and Things

Patents are most commonly justified with reference to their ability to augment the presumptively inefficiently low incentive to produce innovative ideas that exists a legal regime that protects only property rights in tangible resources.[23]

The knowledge required to produce technologically innovative things is costly to produce, in part because of the capital, human and otherwise, that is required to produce it and in part because of the risk, as the production of a valuable idea is rarely certain before the investment is made. Yet, assuming that the investment is made and a valuable idea is produced, competitors are often able to gain access to that idea cheaply and quickly and earn normal profits by selling things embodying those inventive ideas at a price that approaches the things' marginal cost of production. Under these conditions, rational actors will not invest in producing ideas: they see down the road that marginal-cost pricing for goods embodying the ideas will deprive them of the ability to recoup the costs that were sunk into generating the ideas.[24]

Patents alleviate this problem of inefficiently low incentives to generate the ideas that are required to produce new technologies. Critically, however, they do not achieve this goal by what might seem like the most direct means available. They do not treat the newly invented technological ideas *per se* as the resource that is governed by property rights. Thinking about a pill with a previously unknown and unused shape that makes the pill easier to swallow—or talking about it or copying diagrams that represent it—can never constitute infringement of whatever patent issues to the inventor of the newly shaped pill.[25] Rather, patents grant inventors temporally limited rights to exclude others from making, using, selling, offering to sell, or importing a set of things that embody the inventor's new idea(s).[26] If the inventors have generated things that consumers value

[23] [cites for *ex ante* justification] [Note alternatives: incentive to disclose; incentive to commercialize; incentive to efficiently manage (i.e., prospect theory); facilitate transactions in which information is the good exchanged (i.e., deal with Arrow's information paradox)].

[24] [caveat: some incentives exist even w/o patent protection]

[25] Kevin Emerson Collins, *Claims to Information qua Information and a Structural Theory of Section 101*, 4 I/S: A J. OF L. AND POL'Y FOR THE INFO. SOC'Y 11 (2008), *reprinted in* PATENT CLAIMS: JUDICIAL INTERPRETATION AND ANALYSIS (2009) (presenting a structural interpretation of section 101). Recent judicial pronouncements on patent eligibility bring into question whether simply practicing the prior art and then thinking about the knowledge disclosed in a patent specification can infringe a valid patent claim. Prometheus Laboratories, Inc. v. Mayo Collaborative Services, 481 F.3d 1336 (Fed. Cir. 2009). [remanded to Fed Cir to reconsider in light of *Bilski*]

[26] 35 U.S.C. 154(a) (2006).

more than they value existing things, inventors can expect to sell their innovative things at a supra-marginal price during the term of the patent.[27] Inventors may therefore reasonably expect to recoup an investment in the generation of innovative ideas, and they may view investment in innovation as a rational course of action.

Patent rights are therefore intended to function like a chain on a bicycle linking ontologically distinct systems: they grant rights to exclude from sets of things as a means of giving incentives to produce the ideas that are embodied in those things. Despite its centrality in the normative structure that justifies patents, knowledge *per se* remains a spillover of patent protection.[28]

B. Patent Breadth and Depth

When they file their applications, inventors must provide two texts: specifications and claims.[29] Specifications, or disclosures, are descriptions of the invention that usually explain why the invention merits patent protection and that focus on particular examples of the invention. Claims establish the extent of a patentee's rights. The claims of the contemporary patent regime are called peripheral claims because the description establishes the periphery or "outer boundary" of the patentee's rights.[30] Peripheral claims are textual descriptions of the set of things to which patent rights attach.[31]

Although the distinction is not a conventional one, it is helpful to conceptualize the scope of a claim on two distinct dimensions. Patent claims have both *breadth* and *depth*, and a technology infringes only if falls within a claim's scope on both dimensions.[32]

---

[27] [Not all patents generate monopoly power.] Reliance on willingness to pay for things embodying ideas as the measure of social value can be problematic. [elaborate]

[28] [spillover cites]

[29] Technically, the specification includes the claims. []

[30] FABER, MECHANICS OF PATENT CLAIM DRAFTING § 1:2 ("the outer boundaries or periphery of the patentee's claim must be stated").

[31] This set is augmented by a patentee's rights under the DOE. *See infra* [Part III].

[32] Economists employ a distinction similar, but not identical, to claim breadth and depth when they discuss "lagging breadth" and "leading breadth." Ted O'Donoghue, Suzanne Scotchmer, Jacques-Francois Thisse, *Patent Breadth, Patent Life, and the Pace of Technological Progress*, 7 J. ECON. & MANAGEMENT STRATEGY 1 (1998); Ted O'Donoghue, *A Patentability Requirement for Sequential Innovation*, 29 RAND J. ECON. 654 (1998).

Patent breadth is a synchronic concept that collapses the dimension of time.[33] Assuming an omniscient PHOSITA who is aware of all present and future technologies, does the claim language describe—or "read on" in patent lingo—the allegedly infringing technology?[34] Assume that the person who first conceives of a pencil with an attached eraser claims "a pencil with an eraser attached to the end opposite the end used to write."[35] Assume also that a later inventor invents a new eraser/pencil combination in which a highly functional, usable eraser attaches by some ingenious method to the middle of a pencil. There would be debates about the meaning of the words "end" and "opposite," but the improvement likely would not fall within the breadth of the peripheral claim because the middle of a pencil is not its end. The patentee would also argue unsuccessfully that, even if there is no literal infringement of the peripheral claim, there is infringement under the doctrine of equivalents because the allegedly infringing technology is equivalent to the claimed invention.[36] This type of non-infringing technology that lies beyond the breadth of a peripheral claim is the result of "design-around" if the alleged infringer studied the limits on the patentee's interest specified in the claim language and intentionally devised a technology that falls beyond those limits.[37]

Patent breadth is a very familiar concept, but patent depth is less so.[38] Patent depth is a temporal variable that measures post-filing technological progress. It tracks the reach of a claim into technologies that are not invented until after a claim has been filed. At the time a patent is filed, the patent specification constructively discloses a set of newly invented things to the public. For greater specificity, the extent of the disclosure is measured by taking the person having ordinary skill in the art (the PHOSITA) as the target audience.[39] Several patent doctrines tether claim scope to the set of things

---

[33] [cite to synchronic/diachronic distinction] Technically, the patent-breadth determination can be made by fixing the ideational theory of meaning of claim language at the time of filing. *See infra* Part III.B (distinguishing ideational and denotational meaning).

[34] [cite: "read on"] As per the Supreme Court, claim breadth is established through the distinct processes of claim construction—the process of defining the meaning of the words used in a claim—and infringement—the process of determining whether the allegedly infringing technology is described by the claim language, once correctly defined. Markman v. Westview Instruments, Inc., 517 U.S. 370, 384 (1996).

[35] To favor readability and simplicity, this, and all other, hypothetical claims ignore the extensive formalities of claim drafting.

[36] *See generally infra* [Part III.C.4] (discussing the DOE).

[37] [design-around cites]

[38] The lack of familiarity may follow from the fact that, under the commonly held fixation theory of peripheral claims, there is no such thing as claim depth. *See infra* Part III.A (describing the fixation theory).

[39] The set of things disclosed to the PHOSITA by the specification is broader than the preferred embodiments, and has nothing to do with the set of thing-tokens that an inventor actually produces in his or her laboratory. It is, however, limited by the ordinary imagination of the PHOSITA and therefore cannot

constructively disclosed in the specification, and they are addressed in greater detail in conjunction with the "how" question of patent depth in Part III.[40] What is important to understand at this point is only that patent doctrine can restrict patent depth in a manner that is independent of how it curtails patent breadth. Even if the claim language describes the improvement in a synchronic analysis, and thus even if the allegedly infringing technology falls within the breadth of a claim, patent doctrine may still curtail the reach of claim depth into things that were not constructively disclosed to the PHOSITA on the date the patent was filed.[41] For example, working again with the hypothetical patent on the pencil, assume that an improver invents either a nonobvious means of attaching the eraser to the end of the pencil—a new type of glue or mechanical clamp—or a nonobvious improvement in the chemical composition of an eraser. The claim to "a pencil with an eraser attached to the end opposite the end used to write" would seem to have sufficient breadth to encompass either of these improvements. However, the patent must also have depth—patent interest must extend beyond the set of things actually generated by the earlier inventor—for the improvement to infringe the claim. Patent doctrine must allow the set of things that falls within the scope of the claim to extend beyond the set of things disclosed to the PHOSITA at the time of filing, grow over time, and encompass post-filing improvements as they are discovered.[42]

The distinction between patent breadth and depth can be highlighted by revising the graphic conventions through which concepts related to the scope of a peripheral claim are commonly conveyed. Taking a two-dimensional field marked by meaningless X and Y variables in which distinct points represent distinct thing-types,[43] the scope of a peripheral claim is commonly depicted as a circle or amorphous blob that divides those

---

extend into things that are later deemed to be patentable inventions with the patent specification included in the prior art.

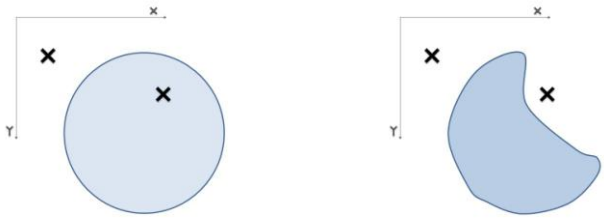
[40] *See infra* text accompanying notes [Parts III.A & C.1.].

[41] Doctrinally, claim depth in isolation may be curtailed in either of two ways: claims may be construed using denotational meaning so that they do not encompass after-arising technologies, *see infra* text accompanying notes [Parts III.B & C.1], or claims that encompass after-arising technologies may be invalidated under the disclosure doctrines, *see infra* [Part III.C.2].

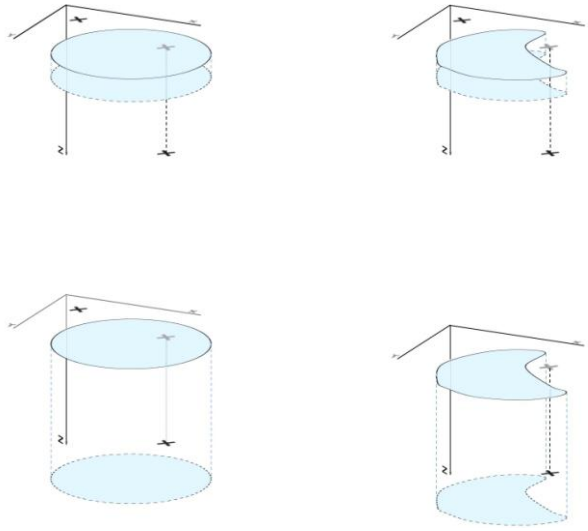
[42] Claim depth can be collapsed into claim breadth by simply stating that the meaning of claim language does not encompass after-arising technology and that claims that are so broad that they encompass too much, or the wrong type, of after-arising technology are invalid. *Cf. supra* note 41 (noting the doctrinal mechanisms for curtailing claim depth). The argument put forward here is not that a breadth-only conception of claim scope is impossible. Rather, the argument is that a breadth-depth model offers a better way to think about patent protection that allows courts to more closely tailor claim scope to fit the needed incentives.

[43] Thing-types, not thing-tokens, are the metric by which claim scope is measured. Collins, *Things and Meanings*, *supra* note [], at 503. Although beyond the scope of this article, even the definition of the thing-types that populate this field is a topic that should be recognized as a source of ambiguity in patent law. *See id.* at 514–36 (defining "thing construction" and demonstrating its importance in patent law).

things that are described by the claim language from those things that are not. The following diagrams depict two different patent claims.



The "X" on the left is a thing that is outside the breadth of both claims, whereas the "X" on the right marks a thing that falls within the scope of one claim but not the other. This conventional diagram depicts only claim breadth. If claim depth is to be added, a third dimension depicting progress in things over time is required. The following figures add claim depth on a Z axis in which time or progress increases as one move down the axis. The top row depicts claims without patent depth, and the bottom row depicts claims with patent depth:



These diagrams presume that "X" on the right—the one that is within the breadth of a claim in the two-dimensional representations on the left—is the product of post-filing progress. It is a thing that was not disclosed to the PHOSITA in the specification at the time the patent was filed. These diagrams reveal that there are two distinct reasons why a claim might not have sufficient temporal reach to encompass an improvement. The upper, left diagram shows a claim in which there is sufficient breadth for infringement, but the lack of patent depth is dispositive of noninfringement. The lower, right diagram shows a claim in which there is sufficient depth for infringement, but the lack of breadth is dispositive of noninfringement. Additionally, the lower, left diagram shows what is necessary for infringement: the claim must have sufficient depth and

breadth. The upper, right diagram shows a situation in which there are redundant reasons for noninfringement—the claim has neither sufficient breadth nor sufficient depth.

C. Cumulative Innovation and Improvement

In the narrative that lies at the core of the justification for intellectual property rights, there is no cumulative innovation. Rights holders are innovators (or their assignees), and infringers are pirates who are free-riders, pure and simple.[44] Yet, it is widely recognized that technological progress is a cumulative endeavor in that the work of earlier generations is a fundamental input into the work of later generations.[45] There may be inventors on both sides of an infringement suit, and intellectual property rights therefore affect the distribution of the rewards of innovation among successive generations of inventors.

At the extremes, there are two ways in which later generations can receive a leg up from earlier generations that are subject to opposing rules of patent protection.[46] On the one hand, the later generation may use the ideas that earlier inventors/patentees disclose in their specifications *qua* ideas—that is, in a purely mental fashion. Because patents meter the consumption of things embodying ideas and not ideas *per se*, these later actors have a use privilege in the earlier-disclosed knowledge.[47] On the other hand, the later generation may use things that the specification constructively disclosed to the PHOSITA and that fall within the breadth of a claim.[48] The later inventor may use an

[44] This narrative is being questioned. Christopher A. Cotropia and Mark A. Lemley, *Copying in Patent Law*, 87 N.C. L. REV. 1421 (2009) (arguing that very few litigated patent cases involve copying).

[45] [sociologists/historians/philosophers of science]

[46] The notion of an earlier inventor giving a later inventor a "leg up" implies that the later inventor actually read and learned from the earlier inventor's patent. However, while copyright does require copying as an element of infringement, Arnstein v. Porter, 154 F.2d 464, 468–69 (2d Cir. 1946), patent protection does not. Therefore, later generations may produce things that are improvements on the things generated by earlier inventors and that infringe earlier patents even if the later generations are unaware of the earlier inventors' works. This assumption simplifies the analysis, but it does not affect this Article's normative and doctrinal conclusions, unless the regime were to adopt an independent inventor defense, *see* Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 275 (2006) (considering the implications of an independent-inventor defense), or in some other way recognize the diversity of fact-specific mechanisms through which later inventors may benefit from earlier inventors, Scotchmer, *supra* note [], at 31 (distinguishing three types of boosts in cumulative innovation: the next innovation cannot be invented w/o the first, the first innovation decreases cost of second, and the first innovation accelerates development of second).

[47] *See supra* Part I.A. Even "pirates" of earlier ideas are off the hook so long as idea used *qua* idea. I can present the information conveyed in a patent specification to large crowds—and even inaccurately attribute the discovery to myself—without running afoul of the patent laws.

[48] Technically, the infringer still uses only the ideas conveyed in a patent, but she uses those ideas not *qua* ideas but to make the very things disclosed to the PHOSITA through the patent specification at the time of filing.

earlier-invented thing (A) as a component in a larger product (A+B)[49] or as a research tool.[50] Here, the very rights that earlier patentees use to fend off pirates—that is, control over the set of claimed things disclosed to the PHOSITA on the date of filing—allow the earlier patentees to control, or at least profit from, the later generations' work.[51]

As the term is used in this Article, *improvers* are second-generation inventors who lie in between these two extremes.[52] Improvers do not make the things disclosed to the PHOSITA in earlier patents. They make *improvements*—new things that embody post-filing progress. They use the earlier inventor's ideas as inputs into the process of inventing new things—things that often that strongly resemble the things actually generated by the earlier inventor and that are usually economic substitutes for those things.[53] Given this definition of improvements, earlier patents must have depth if improvements are to infringe them. By definition, an improvement is not a member of the set of things that was disclosed to the PHOSITA by an earlier patent's specifications. Improvements are new things that entail technological advances over the prior art which, at the time of the improvement, includes the disclosure of the earlier patent.

This definition of an improvement is relatively narrow in one way that is important to highlight. In an every-day, loose sense of the word, later inventors who develop new things that are complementary to earlier-patented things are commonly called improvers. For example, assume that an earlier inventor patents chemical A which is useful as a cleaner. A later inventor may invent chemical B, an additive that does not change the chemical structure of chemical A but that increases the efficacy of chemical A as a cleaner when it is mixed with chemical A. In an every-day sense, the later inventor has created an improvement, as A+B cleans better than A does. In the more technical sense in which the concept of an improvement is used in this Article, however, there is no improvement. The combination A+B infringes the earlier claim to A not because A+B is an improvement but because the A in A+B is the same old A that the earlier inventor

[49] For example, smaller chips (A) allow later inventors to design innovative, light-weight devices (B) that contain the chips (A+B). [cites to component-industry controversies]

[50] A classic example of a research tool is a microscope that helps later researchers and inventors to achieve further progress. [note on research tool controversies (Eisenberg & Strandburg), including the distinction between researching on and researching with]

[51] Research tools and components raise distinct policy concerns for patent law because they often correlate to the difference between a one-to-many and a many-to-one relationship between the inventions in earlier and later generations. SUZANNE SCOTCHMER, INNOVATION AND INCENTIVES 131–32 (2004).

[52] A distinct scenario that also lies in between these extremes involves reverse engineering. Later users may make and/or use a thing actually disclosed in the patent specification in order to learn knowledge about the invention and to use it *qua* knowledge. *See generally* Pamela Samuelson & Suzanne Scotchmer, *The Law and Economics of Reverse Engineering*, 111 YALE L.J. 1575 (2002).

[53] Improvements can sometimes be complements to the things disclosed and claimed in an earlier patent. *See infra* note [].

disclosed to the PHOSITA in a patent specification. The later inventor has only used A as a component in combination with a distinct, complementary thing, B. To be an improver, the later inventor must produce a new thing that is distinct from the things disclosed to the PHOSITA in the earlier patent and does not simply contain those earlier-disclosed things as a component. She must produce an A'.[54] The following two diagrams illustrate the difference between the A+B after-arising component scenario (on the left) and the A' improvement scenario (on the right). The critical difference is that no patent depth is required for an earlier patentee to control A+B, but patent depth is required for the earlier patentee to control the A' improvement:

As a corollary to the distinction between improvements and after-arising components, it is important to recognize a distinction between two types of blocking patents, one of which can exist without patent depth but the other of which cannot. Blocking patents arise when two or more patentees possess a right to exclude others—including each other—from making, using, or selling any given technology.[55] *Distinct-thing* blocking patents describe two distinct complementary goods that consumers want to consume together. If there is a first patent that encompasses A and a second patent that encompasses B, the technology of A+B is subject to distinct-thing blocking patents. Neither of the patents needs to have any depth at all for distinct-thing blocking patents to develop, even if B is not developed until many years after A. In contrast, *same-thing* blocking patents both describe the self-same thing. If there is a first patent that claims A, a later patent that claims A', and A' does not simply incorporate A as a component, A' is subject to same-thing blocking patents if and only if the patent on A has sufficient depth to encompass A'.[56] Because an improvement is (as per this Article's narrow definition)

[54] Economically speaking, the need for patent depth to make A' (an improvement) infringe but the lack of any need for patent depth to make A+B (a later-developed technology that includes the earlier-disclosed thing as a component) infringe is difficult to defend. The distinction is a bias that is introduced into patent protection by the thing-centric nature of peripheral claims. *Cf.* Collins, *Things and Meanings*, *supra* note [], at 514–36 (discussing "thing construction"). In fact, this Article's conclusion that literal claim scope should routinely encompass orthogonal improvements would mitigate the formalistic distinction between later inventors who produce B and later improvers who produce A'. *See infra* Part II.D.3 (arguing that orthogonal improvements should be treated like bundles of earlier-patented and later-developed distinct things).

[55] [cites: blocking patents]

[56] *See generally* Collins, *supra* note [], at [] (distinguishing same-thing and distinct-thing blocking patents).

by definition an A', patent depth is a prerequisite for a later improvement to infringe an earlier patent and for same-thing blocking patents to develop.

As a practical matter, infringement cases involving the temporal reach of patent scope into improvements can be divided into two sets. Some improvements will fall beyond the breadth of a claim, in which case the possible absence of claim depth is irrelevant because being beyond the breadth of a claim is a sufficient reason for non-infringement.[57] Other improvements, however, will fall within the breadth of a claim, and the issue of claim depth will be dispositive.[58] As a doctrinal matter in Part III (addressing the "how" question), this distinction is critical, as courts use different policy levers to adjust claim breadth and depth.[59] As a normative matter in Part II (addressing the "when" question), however, the difference does not loom as large. The same economic arguments that suggest claims should be sufficiently deep to encompass a particular type of improvement can often be mustered to suggest that claims should be sufficiently broad to encompass that type of improvement.[60] Part II therefore addresses the temporal reach of patent scope into improvement as a generic concept, but Part III returns to the distinction between patent breadth and depth.

II. RETHINKING THE "WHEN" QUESTION

The conventional wisdom concerning the conditions under which an earlier patent should have sufficient temporal reach to encompass an improvement has developed with a particular type of improvement in mind: the classic improvement. The conventional wisdom has completely ignored a distinct type of empirically common improvement—the orthogonal improvement—that results from advances that are cumulative in a different way. This oversight makes the conventional wisdom on the "when" question of the temporal reach of patent scope incomplete. The conventional wisdom suggests that the optimal temporal reach presents a highly contested issue. However, the optimal temporal reach for orthogonal improvements should present an easy case. Earlier patents should routinely have sufficient reach to encompass orthogonal improvements.[61] Part II.A describes and illustrates the classic improvement. Part II.B briefly summarizes the conventional wisdom on the "when" question of temporal reach for improvements—

[57] *See supra* text accompanying note [] (discussing the pencil with an eraser attached to the middle).

[58] *See supra* text accompanying note [] (discussing pencils with erasers made from after-arising materials and with after-arising mechanisms of attachment).

[59] *See generally infra* Part III.C.

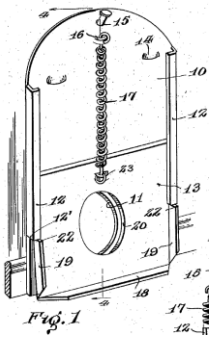
[60] Empirically, however, there seems to be a correlation between cases involving orthogonal improvements and cases in which patent depth, rather than patent breadth, is the dimension that would most likely limit the temporal reach of patent scope. [bald assertion; flush out]

[61] In other words, the distinction between classic and orthogonal improvements should be added as a new factor to the conventional wisdom on the "when" question of patent depth.

wisdom that has developed with the classic improvement as its implicit target. Part II.C describes and illustrates the overlooked improvements. Part II.D develops a conceptual framework that both descriptively captures the relevant distinction between classic and overlooked improvements and normatively explains why temporal reach should routinely encompass overlooked improvements even as temporal reach for classic improvements remains a contested issue. Part II.E reflects on the parallels between the classic/overlooked distinction developed in this Article and the cumulative/discrete distinction articulated by Merges and Nelson in their seminal article on patent scope and improvement.[62]

A. Classic Improvements

The easiest way to introduce the classic improvement is with a hypothetical that uses a simple technology. An initial inventor, Abbey, is the first to invent a spring-loaded mousetrap. Before Abbey's invention, the state-of-the-art mousetrap was an upside-down box over a piece of cheese with a short stick holding one side of the box above the floor. The person attempting to catch the mouse would tie a string to the stick, wait nearby, and pull when a mouse went under the box. Abbey invents the first spring-loaded mousetrap: a device that stores energy in a spring and that uses the jostling motion caused by the presence of a mouse to release kinetic energy, trapping or killing the mouse. The mousetrap that Abbey actually produces and discloses is illustrated in Figure X:

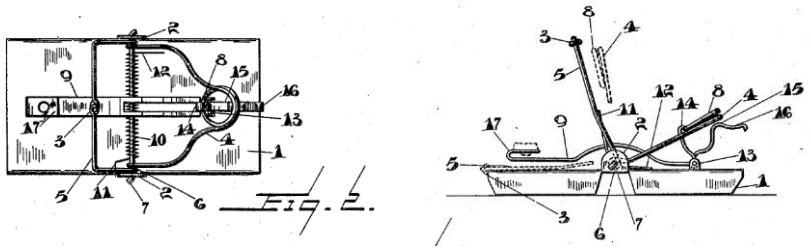


There are two plates, each having a hole in the center and one being able to slide vertically in relation to the other. Cheese is placed in a box with Abbey's mousetrap forming one side of that box, or the mousetrap is placed over a mouse hole in the wall. The spring must be stretched from its resting position for the two holes to align so that a mouse can enter the box. Although there is a stop mechanism that can hold the holes in alignment and keep the spring in tension, the jostling motion caused by a mouse who passes through the aligned holes destabilizes the stop mechanism, allows the spring to shift one plate with respect to the other, and traps or kills the mouse. Abbey files a patent that broadly claims "a device that stores potential energy in a spring, wherein the

[62] Merges & Nelson, *supra* note [].

potential energy is converted to kinetic energy that traps or kills the mouse when a mouse is present."[63]

After Abbey files her patent, Bernard invents an improved spring-loaded mousetrap that is, more or less, the standard mousetrap design that one can still buy in the local hardware store today. General familiarity with such mousetraps is presumed:



Where Abbey's trap keeps the spring in tension longitudinally, Bernard's trap places a torsional force on the spring. Where Abbey's trap involved sliding plates, Bernard's trap has a wire moving in an arc in relation to the base. Bernard's mousetrap is a patentable improvement over the disclosure of Abbey's patent, although Bernard may or may not seek patent protection.

Initially, it is important to note that the Abbey-Bernard hypothetical illustrates that there are two distinct technological advances made at different points in time whenever the question of the temporal reach of patent scope for improvements is raised.[64] First, there is the advance in technical knowledge that justifies the issuance of the earlier patent whose temporal reach is at issue. The things described by a patent claim embody an advance by definition as a doctrinal matter: it is the advance that explains why the things described by a claim satisfy the validity requirements of patent law. Thus, all the things encompassed within an earlier novel and nonobvious claim embody an advance. Second, every allegedly infringing improvement embodies an advance produced by the improver that occurs after the time the patent is filed. If it didn't, it would be an improvement over the technology disclosed in the earlier patent, and patent depth would not be required for infringement. The improver may or may not seek patent protection for her improvement. Because the issue at hand is the temporal reach of the earlier-issued patent, however, whether the improver actually seeks patent protection is irrelevant.

[63] For the sake of readability and simplicity, this claim ignores the well-established conventions of claim drafting.

[64] An advance is the idea or bit of technical knowledge generated by anyone who pushes technological progress forward. There is no specific kind of idea that constitutes an advance. An advance is dependent upon the nature of the knowledge an inventor contributes to the storehouse of knowledge.

The Abbey-Bernard hypothetical strongly resembles many of historical examples of earlier patents and allegedly infringing later improvements.[65] The family resemblance follows from two facts about the way in which the later advance the earlier advance. First, the work of an earlier inventor builds on the work of a later inventor in the sense that the later inventor must stand on the shoulders of the earlier inventor in order to be in a position to make her contribution to technological progress. But for Abbey's idea of a spring-loaded mousetrap in general, Bernard would not have been able to invent anything at all (or, more accurately, invent anything at all at the time at which he did) because Bernard invented a particular type of spring-loaded mousetrap.[66] But for Abbey's shoulders, Bernard would not have been in a position to make the advance that he did when he did it. Second, however, Bernard's later technical insight does not simply add to or compound with Abbey's technical earlier insight. Bernard refines Abbey's insight in the sense that his advance in part replaces, supplants, or revises Abbey's advance. Because of this revision, Bernard's spring-loaded mousetrap is in a meaningful way no longer entirely Abbey's mousetrap.

It is this mixed relationship between the earlier and later advances—their part-compounding and part-supplanting nature—that defines the classic improvement. The hybrid is possible because Abbey's patented technology embodies a technological advance that can be described at different levels of generality. At a high level of generality, the advance embodied in Abbey's mousetrap is the very idea of a spring-loaded mousetrap itself—i.e., the idea of a mousetrap that can store potential energy in a spring and automatically unleash mouse-trapping kinetic energy in response to the presence of a mouse. At a lower level of generality, the advance embodied in Abbey's mousetrap is the idea of storing potential energy in an elongated spring and releasing it as kinetic energy in the form of one plate with a hole that slides in relation to another. The after-arising advance that gives rise to Bernard's improvement is a new mechanism for storing potential energy and using kinetic energy to catch a mouse. Bernard's mousetrap therefore compounds with or adds to Abbey's most general advance in the improvement. The improvement fully embodies both Abbey's general idea of a spring-loaded device actuated by the presence of a mouse and Bernard's specific idea of a device that stores energy in a torsional force and releases it in the form of a wire traveling in an arc. However, Bernard's advance supersedes Abbey's more specific advance of connecting an elongated spring to slidable plates in the improvement. By storing potential energy in a torsional force, Bernard's mousetrap no longer stores energy in a longitudinal force. By using a wire moving in an arc, Bernard's mousetrap no longer uses sliding plates with holes. Any single, indivisible mousetrap will embody either Abbey's specific idea or

[65] *See, e.g.*, *supra* text accompanying notes [] (discussing the Selden and Wright Brothers' patents). *See also* James Bessen and Eric Maskin, *Sequential Innovation, Patents, and Imitation*, 40 RAND J. ECON. 611, 612 (2009) ("By 'sequential,' we mean that each successive invention builds on the preceding one, in the way that the Lotus 1-2-3 spreadsheet built on VisiCalc, and Microsoft's Excel built on Lotus."). [other cites]

[66] *See supra* note [] (noting that patent infringement presumes appropriation, not independent invention).

Bernard's specific idea, but not both.[67] In sum, in a classic-improvement scenario, the after-arising advance compounds the earlier advance when the latter is framed at a high level of generality, but it supersedes the earlier advance framed at a low level of generality.

B. Some Conventional Wisdom Concerning the "When" Question

The conventional wisdom concerning the "when" question of the temporal reach of patent scope is complicated, to say the least. In general, the problem is often framed as the problem of distributing the incentive to invent among different generations of inventors.[68] There is also general agreement that the limited duration of patents is one feature of the patent regime that helps to allocate rights between different generations of inventors.[69] Yet, beyond these few settled points, one set of economic assumptions and theoretical commitments seems to justify the position that patents should rarely, if ever, have sufficient temporal reach to encompass classic improvements, but a different set leads to the opposite conclusion that patents should have sufficient temporal reach to routinely encompass classic improvements. Furthermore, there is a large space in the middle for suggesting that the availability of temporal reach for classic improvements should be a case-by-case determination that is highly contingent on the particular facts at hand. This Part offers a brief overview of the principal factors that are conventionally viewed as making the answer to the "when" question of the temporal reach of patent scope a highly contested issue.

1. The Pioneering Nature of the Earlier Invention

Pioneer theory gives special consideration to path-breaking inventions in determining the reach of an earlier inventor's patent into improvements.[70] Path-breaking

[67] It is possible for a mousetrap to embody both in the sense that Abbey's mousetrap and Bernard's mousetrap can be glued together to form a double-wide mousetrap with two trigger mechanisms. This physical aggregation of the earlier- and later-invented things, however, does not raise the question of patent depth. *See supra* text accompanying notes [Part I.C] (noting that A+B infringes a patent on A even if the patent has no depth).

[68] Scotchmer, *supra* note [], at 30 ("The challenge [when dealing with cumulative innovation] is to reward early innovators fully for the technological foundation they provide to later innovators, but to reward later innovators adequately for their improvements and new products as well."). *But see infra* [Part II.B.5] (discussing the prospect theory of patent law).

[69] [cites] However, if patents lack depth, then the effective life of a patent is much shorter than its legal term because design-around will usually be trivial. O'Donoghue, Scotchmer, & Thisse, *supra* note [], at []. For a strong argument that limited term is the only restriction needed to craft patent protection in a manner that efficiently fosters cumulative innovation, see Richard Epstein, *The Disintegration of Intellectual Property? A Classical Liberal Response to a Premature Obituary*, 62 STAN. L. REV. 455 (2010). It is interesting to note that Epstein is silent on the topic of how patent scope in general—whether for improvements or not—should be determined.

[70] John R. Thomas, *The Question Concerning Patent Law and Pioneer Inventions*, 10 HIGH TECH. L.J. 35 (Fed. Cir. 1987). The contemporary vitality of pioneer theory is suspect as a doctrinal matter.

inventions are inventions in which the contribution of an inventor at a high level of generality is particularly important. However, the specifics of path-breaking inventions are often not yet well developed at the time a patent is filed. This means that several generations of post-filing inventions are very likely to follow a path-breaking invention as different paths are explored, the invention is refined, and commercial products become available. Pioneer theory presumes that path-breaking inventions are extremely valuable to society, so the patent regime should give them liberal encouragement. It also recognizes that a claim without much of a reach into later-developed improvements would provide little reward to the path-breaking inventor because so many of the products for which consumers will be willing to pay will be improvements on the disclosed technology. Pioneer theory therefore suggests that the patents of pioneer inventor patents should be endowed with additional temporal reach, allowing them to internalize more of the value of the others improvements.[71]

2. The Pattern of Technical Advance in the Industry

Robert Merges and Richard Nelson posit that the temporal reach of patent scope should turn in part on the pattern of technical advance that exists in an industry.[72] Some industries are "discrete" innovation industries in which an invention "does not point the way to wide ranging subsequent technical advances."[73] Other industries are "cumulative" innovation industries in which "today's advances build on and interact with many other features of existing technology."[74] Merges and Nelson are concerned about the impact of broad, earlier patents on the pace of later improvements in cumulative innovation industries but not in discrete innovation industries. "Particularly when the technology is in its early stages, the grant of a broad-gauged pioneer patent to one party" in a cumulative innovation industry "may preclude other inventors from making use of the inventions without infringing the original patent."[75] The gist of the argument is two-fold. First, permitting competition for follow-on innovation is better than allowing a single firm to control the direction of follow-on innovation.[76] Second, the patent licensing that would be necessary to facilitate a competitive market for follow-on innovation if there is

[71] *Cf.* Howard F. Chang, *Patent Scope, Antitrust Policy, and Cumulative Innovation*, 26 RAND J. ECON. 34 (1995) (elaborating an economic model that supports broad protection for inventions "with very little stand-alone value relative to the improvements that they may inspire").

[72] Merges & Nelson, *supra* note [], at 880−908.

[73] *Id.* at 880.

[74] *Id.* at 881. Merges and Nelson also separate out chemical and science-based industries for distinct treatment. *Id.* at 897–908.

[75] *Id.* at 882.

[76] *Id.* at 908 ("[M]ultiple and competitive sources of invention are socially preferable to a structure where there is only on or a few sources.").

a broad and deep patent should not be presumed and may not occur. Therefore, courts should, at the margin, construe patents not to encompass improvements in industries that are characterized by cumulative innovation.[77] Interestingly, where pioneer theory suggests that the reach of patent scope into improvement should be relatively liberal in cumulative innovation industries,[78] Merges and Nelson suggest precisely the opposite.

3. The Value of the Improvement

A similar lack of faith in the market for patent licenses motivates the proposal that particularly valuable improvements should lie beyond the temporal reach of earlier patents. In the actual world that is rife with strategic behavior, imperfect information, and transaction costs, there are good reasons to fear "bargaining breakdown" when a radical improvement falls squarely within the scope of an earlier claim.[79]

4. The Magnitude of the Desired Patent Incentives

Economists interested in cumulative innovation sometimes work with a *quality ladder* as a stylized model of improvement.[80] The quality-ladder model assumes that firms create successively better products, each improving on the previous one. No innovator is identified as the first to create a product, and no innovator is secure in her belief that she will be the last to improve the product. Furthermore, each successive innovator is assumed to produce an improvement of that is of higher quality than the previous one and that is therefore more desired by consumers.[81]

The optimal temporal reach of a patent on a quality ladder depends on the desired magnitude of the incentive to innovate to be created by the patent regime. The longer the

---

[77] *Id.* at 843–44. Merges and Nelson "take it as axiomatic that some degree of patent protection [for technologies in cumulative innovation industries] is necessary and desirable." *Id.* at 885. *But see* James Bessen and Eric Maskin, *Sequential Innovation, Patents, and Imitation*, 40 RAND J. ECON. 611, 612 (2009) (arguing that the absence of patent protection may be preferable to patent protection in industries that are characterized by "sequential" and "complementary" innovation—roughly industries that resemble the cumulative innovation industries discussed by Merges and Nelson).

[78] *See supra* Part II.B.1 (making this distinction).

[79] Robert Merges, *Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents*, 62 TENN. L. REV. 75 (1994). *See also* Lemley, *Improvement*, *supra* note [], at 1061–65; Merges & Nelson, *supra* note [], at 909 (noting that courts "should consider . . . the importance of the advance represented in the *accused* device" in the infringement determination). In the contemporary patent regime, the value of an improvement is relevant only under the rarely used reverse DOE, not under the doctrines that affect literal infringement or infringement under the DOE. Lemley, *supra*, at 1006.

[80] SCOTCHMER, *supra* note [], at 149–52.

[81] The quality ladder model is a highly stylized model of patent protection because the value of an allegedly infringing technology is normally not relevant to the infringement determination. *See supra* note 79.

reach of patents into after-arising technology, the larger the supra-competitive profits that is created. Assume at first that an improvement in quality of Δ or greater lies beyond the reach of an earlier inventor's rights.[82] Each improver on the ladder must compete in the market with an earlier product which is Δ less valuable and, once an improvement is made, with a later product with is Δ more valuable. Second, assume that the temporal reach extends to any improvement with a quality difference of less than 2Δ. The inventor must share her profits in any given time period through a licensing agreement (at first with an earlier patentee and then with a later patentee), but this is balanced out by the fact that her patent covers the highest quality product over two, rather than one, time period. The difference maker is that the quality difference between the highest-quality, patented product and its closest rivals is twice as great, so the patentee's per-period profits will be larger. For this reason, the temporal reach of a patent should be greater when the optimal strength of the patent regime's incentives to innovate is higher.[83]

5. The Import of Patent Law's Prospect Function

Pioneer theory identifies a subset of inventions—those that open up new terrain for improvers—and affords them additional temporal reach in order to optimize the incentives for innovation.[84] The "prospect theory" of patent law suggests that patent law should achieve these very ends for very different reasons: to promote efficient development of nascent technologies into the improved products that are valued by consumers.[85]

[to add: recap of prospect theory]

C. Overlooked Improvements

There is an entire category of empirically common improvements that are not classic improvements. Again, an example is instructive. Taking Abbey's invention of the spring-loaded mousetrap as the earlier-patented invention,[86] assume that Bob, too, produces an improved mousetrap. Bob invents a nonobvious metal alloy which makes

[82] Assume also that such improvements occur before the expiration of the earlier patent and that the earlier and later inventions are made by different people.

[83] A larger patent-induced incentive is preferable if the costs of innovation are high and there are few non-patent means for inventors to internalize some of the value of their inventions. [Burk & Lemley on industry specificity & policy levers]

[84] *See supra* Part II.B.1.

[85] Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977).

[86] *See supra* text accompanying notes [Part II.A].

cheaper, better springs.[87] Bob manufactures sliding-plate mousetraps that follow the precise arrangement of mechanical parts that Abbey discloses in her specification, except that he makes his springs out of his after-arising alloy. Bob's mousetrap is clearly an improvement. It embodies a nonobvious advance in relation to the state of the art at the time Abbey files her patent (with Abbey's disclosure as part of that art); it is an after-arising thing that was not disclosed to the PHOSITA in Amy's patent under the enablement and/or written description doctrines; and Amy's patent must have depth in order to encompass it.[88] Yet, it differs from a classic improvement in two ways.

First, the improver does not need to stand on the shoulders of earlier patentees as a precondition of arriving at the after-arising advance that gives rise to the improvement. Bernard, the classic improver, benefits from a boost from Abbey in order to be in a position to achieve his advance.[89] Bob does not. The technical barrier confronting Bob would be the same whether Bob produces his advance in metallurgy before or after Abbey makes her advance in mousetrap technology.[90] As idea-generators who contribute to technological progress, Bob and Abbey stand side by side, not one on the shoulders of the other. However, the improvement-as-thing exists only because of the accumulation of the two advances—the improvement can be held aloft only by a concerted effort of Bob and Abbey, if you will, even though Bob and Abbey are standing side by side. The improvement is the result of cumulative innovation only in the weak sense that two advances both must occur for the improvement to be produced, but neither one of the advances in knowledge builds on the other.

Second, the earlier and later advances wound up in the story of the improvement are both fully embodied in the improvement. Bob's after-arising advance compounds with, and does not supersede, Abbey's earlier-patented advance, regardless of the level of generality at which Abbey's advance is framed. This pattern differentiates Bob's improvement from Bernard's improvement because Bernard's advance in part superseded

---

[87] Any one of a number of different advances could underlie Bob's discovery. He may have been the first to conceive of a molecule with a particular chemical structure or, more likely, he may have been the first person to figure out how to make a long-desired compound. This distinction is of importance in determining optimal patent depth in some contexts, but it is irrelevant in this hypothetical.

[88] *See supra* Parts I.B & C.

[89] *See supra* text accompanying notes [Part II.A].

[90] Interestingly, however, if Bob were to invent first, Bob's patent would not need to be construed with any depth for a mousetrap made following Abbey's design and using a spring made from Bob's alloy to infringe. Bob would likely claim a new molecule—an extremely narrowly framed thing that is literally present in the improved mousetrap. This asymmetry—the need for Abbey's patent to have depth to encompass the improvement but the lack of a need for Bob's patent to have depth to encompass the same improvement if Bob's patent were to have come first—demonstrates the bias that the nature of things introduces into the patent regime. *See generally* Collins, *Meanings and Things*, *supra* note [], at 514–36 (discussing the importance of thing construction in patent law).

Abbey's advance.[91] Bob's improvement embodies his after-arising advance (the idea of the new alloy) as well as both Abbey's earlier-patented general advance (the idea of a self-actuating, spring-loaded mousetrap) and her earlier-patented specific advance (the idea of an elongated spring connected to a sliding plate). Bob has improved on features of Abbey's patented technology, but those improved features that have nothing to do with the features that embody the ideas that Abbey was responsible for generating. Bob did not make an advance in the arrangement of the mechanical components in a mousetrap design; Abbey did not make an advance in any way related to the molecular structure of the components. Unlike in the classic improvement, the properties of the mousetrap-as-thing that embody Abbey's earlier-patented advance are in no way displaced by properties that embody the improver's after-arising advance. For this reason, there is an intuitively meaningful sense in which the mousetrap produced by Bob is still also fully Abbey's mousetrap, just made out of a different material.[92]

D. The "When" Question for Overlooked Improvements

This Part argues that the conventional wisdom addressing the "when" question of the temporal reach of patent scope for improvement does not pertain to overlooked improvements. Unlike classic improvements, which are widely viewed as cases in which the temporal reach of claim scope is highly contested, overlooked improvements should present easy cases. Earlier-issued patents should routinely have sufficient temporal reach to encompass them.

Part II.D.1 develops the conceptual framework required to describe the distinction between the two types of improvement. It argues that invention-space—the metaphorical space in which progress in the technological things claimed by patents can be mapped—must be recognized as a multidimensional space. Classic improvements involve a later advance that refines an earlier advance that implicated the same dimension of invention-space, whereas the overlooked improvements involve successive advances along two independent dimensions of invention-space. For this reason, the overlooked improvements are called orthogonal improvements.

The remaining two parts articulate the normative justifications for having different rules governing the temporal reach of claim scope for classic and orthogonal

[91] *See supra* text accompanying notes [Part II.A] (noting how Bernard's advance is part-compounding and part-superseding).

[92] If Abbey's advance were characterized as the idea of making a spring-loaded mousetrap out of then-existing metals, then Bob's advance could be viewed as a part-superseding advance. Thus, the distinction between classic and overlooked improvements may be one of degree rather than kind. Even if this is true, the distinction retains its importance as a valuable descriptive and normative concept. However, the characterization of Abbey's advance as the idea of making a spring-loaded mousetrap out of then-existing materials, as it elides the things that Abbey invented with the contribution to the storehouse of knowledge that she made. To describe the thing that Abbey made, it is necessary to note that she worked with existing metals. However, her use of existing metals is in no way necessary to describe her advance, i.e., her marginal contribution to the progress of technical ideas.

improvements. Part II.D.2 invokes the proportionality principle—the principle that the private value of patent rights should be roughly proportional to the social value of the ideas generated by inventors. Part III.D.3 argues that orthogonal improvements are the embodiment of two technological advances that are economic complements in the same, indivisible thing. To extend the economic treatment given to the successive producers of complementary things (like hammers and nails) to the successive producers of complementary properties of things (like designs for mechanical devices and the materiality of mechanical devices), patents must routinely have sufficient temporal reach to encompass orthogonal improvements.

1. Multidimensional Invention-space and Orthogonal Improvements

One assumption that underlies discussions of the common wisdom on the "when" question of the optimal temporal reach of patent scope is that the improvement of a thing always registers on a linear scale. This assumption is made expressly in the quality ladder model of improvement.[93] It is implicit in the notion that the social value or the importance of the improvement in relation to the earlier-patented technology should influence a patent's temporal reach, as social value is conventionally measured only by its magnitude.[94] Pioneer theory and prospect theory recognize that a path-breaking invention may be developed in a number of different directions, yet the direction taken remains irrelevant to the issue of the patent's temporal reach. There is no mention in either theory of the notion that developments in some directions should be treated differently than developments in other directions.[95] The distinction between discrete and cumulative technical advance on which Merges and Nelson focus is a distinction between kinds, but it is most commonly framed as a variable that is about the nature of the industry in which an invention occurs and as exogenous to the nature of the particular improvement at issue.[96] What remains important in each case of improvement is the magnitude of the progress achieved by the earlier patentee and the magnitude of the progress achieved by the improver.[97]

The distinction between classic and orthogonal improvements challenges the assumption that improvement is a phenomenon that can be measured in a linear fashion

[93] *See supra* Part II.B.4.

[94] *See supra* Part II.B.3.

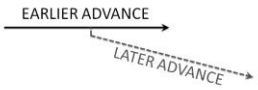
[95] *See supra* Part II.B.1 & 5. These theories measure improvement on polar coordinates in which the angular variable is meaningless.

[96] *See supra* Part II.B.2. *But cf. infra* Part II.E (arguing that the distinction between cumulative and discrete technical advances can be roughly correlated with the distinction between classic and orthogonal improvements).

[97] Lemley, *Improvement*, *supra* note [], at 1070 (arguing that rights to improvements should turn on "the *relative* importance of the original invention and the improvement as the critical factor").

by magnitude alone. Different types of improvement are qualitatively different in a way that should make a difference. The difference stems from the relationship between the two technical advances that are wound up in the question of the temporal reach of patent scope into improvements, namely the earlier advance that justifies the issuance of the patent and the later advance that gives rise to the improvement.[98] These two advances can be cumulative in two distinct ways. To explain this difference, the metaphorical space in which the technological progress of things can be mapped—call it *invention-space*—must be acknowledged as multidimensional. In a multidimensional invention-space, not all successive advances in technological knowledge that lead to improvements need to be cumulative in the same manner.

Classic improvements arise when the earlier and later advances implicate progress on a common dimension of invention-space, with the later advances refining earlier ones. Graphically presented, the advances that lead to a classic improvement look like this:

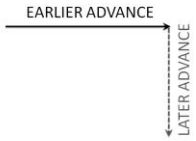


The later advance builds on the earlier advance. But for the earlier advance, the inventor who conceived the later advance would not have had the jumping-off point that she did. However, the later advance also supplants part of the value of the earlier advance. The technical advance that underlies the earlier patent is in part rendered outmoded or obsolete.[99]

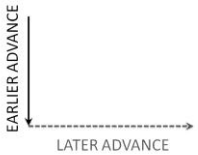
In contrast, overlooked improvements involve successive advances on orthogonal dimensions of invention-space. For this reason, this Article refers to the overlooked improvements as *orthogonal* improvements. The after-arising advance that gives rise to the improvement operates along a dimension of invention-space that is orthogonal to—and thus independent of—the dimension that charts the advance that justifies the earlier-issued patent. Graphically presented, the advances that lead to an orthogonal improvement look like this:

[98] *See supra* text accompanying notes [Part II.A] (highlighting the two distinct advances wound up in the question of patent depth for improvements).

[99] Technical superiority is not needed for an improvement patent. Improvement may offer alternative but not render the earlier patent outmoded. [cite: utility does not require a better-than-the-prior-art standard]



The improvement-as-thing results from the accumulation of the two advances. However, the later advance does not build on the earlier advance, nor does the later advance supplant any of the value of the first advance. Because each advance stands on its own technological bottom, it is entirely possible to imagine getting to the same point—that is, producing the same technological thing—with the two advances occurring in the opposite order, with the later advance moving earlier in time and the earlier advance moving later in time:



This inversion is not possible in a classic improvement, because the earlier inventor's advance is a foundation for the later inventor's advance.[100]

[100] One can imagine a possible world in which Bernard's mousetrap is produced before Abbey's. Graphically, it would look something like this:

In this possible world, however, Bernard must make a different technological advance than he did in the actual world. He must do what Abbey did in the actual hypothetical, namely generate the general idea of the spring-loaded mousetrap. This is why the advances in a classic improvement cannot be inverted.

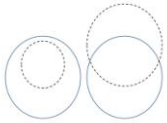
What does it mean as a practical matter for two successive advances to be charted on orthogonal dimensions in invention-space? How does one identify and define any given dimension of invention-space? In part, as already discussed, advances are orthogonal because they rest on their own technological bottoms and neither one advance nor the other must come first.[101] In other words, the inventors stand side-by-side, not foot-on-shoulder.[102]

This independence of the prior technological knowledge required for each of the advances, however, is not a sufficient criterion for identifying orthogonal improvements.[103] Additionally, the properties that manifest the successive advances in the improvement must be capable of some degree of independent variation. Advances embodied in the two properties implicate progress on distinct dimensions of invention-space only when there is this type of functional independence between the two properties.

Things are extremely complex entities. They possess, or perhaps are comprised of, an enormous array of distinct properties.[104] Properties are critical entities in patent law. To correct a market failure for innovative ideas, patents grant inventors rights to exclude others from making, using, or selling things that embody newly discovered

---

[101] Because the later inventor's invention that gives rise to an orthogonal improvement is often in a different technological field, orthogonal improvements often implicate a different type of *same-thing* blocking patent (assuming that the later inventor also seeks patent protection). *See supra* text accompanying notes [Part I.C] (distinguishing same-thing and distinct-thing blocking patents). Orthogonal improvements are likely to involve overlapping patents (on the right), whereas classic improvements are more likely to involve nested patents (on the left).



[102] Tim Holbrook has argued that the DOE should more readily encompass after-arising technologies when the later advance occurs in a field of endeavor that is different from the field of endeavor of the patent. Timothy R. Holbrook, *Equivalency and patent Law's Possession Paradox*, 23 HARV. J.L. & TECH. 1, 7 (2009) (arguing that, under the DOE, "protection for later-developed technologies is appropriate when the advance took place outside of both the inventor's field of technology and those fields of which the inventor should have been aware."); *see also id.* at 37–40. Holbrook justifies this based on a "fairness principle." *Id.* at 7. This Article argues that the distinction is a relevant factor—but not the only factor—when determining the conditions under which literal claim scope should encompass after-arising technology. It also argues that the distinction can be explained as a matter of efficiency, not only fairness. *See infra* Part II.D.2 & 3.

[103] [provide an example of advances that derive from distinct technological bases and yet that do not yield orthogonal improvements]

[104] "Properties" is used here in the sense that philosophers of things use the term. [cites to properties in philosophy of things] No conceptual relation to "property" as a legal term is intended.

technological advances, i.e., ideas.[105] What it means for a thing to embody and advance/idea is nothing more than what it means for a thing to possess certain properties. A mousetrap embodies Abbey's newly discovered specific idea in part because it has the property of *being made of sliding plates*,[106] a mousetrap embodies Bernard's newly discovered specific idea in part because it has the property *being made of a wire that can move in an arc in relation to a fixed base*,[107] and a mousetrap embodies Bob's newly discovered idea because it has the property *having the molecular structure of the after-arising alloy*.[108] Critically—here is where we reach the crux of the matter—not all of a thing's properties need to be equally implicated in the group of properties that embody an advance in a thing. In some instances, there are two distinct properties, or sets of properties, of a thing, each of which is capable of being altered or changed within certain parameters without having a significant impact on the other. When the successive advances wound up in stories of improvement and a patent's temporal reach are manifest in properties of things that are independent variables in this manner, the advances can chart progress on orthogonal dimensions of invention-space.

There is no single, catch-all explanation for why two properties are independent variables and thus why successive advances that are embodied in things as those two properties map progress on orthogonal dimensions of invention-space. In fact, two clear examples of orthogonal improvements suggest two different underlying reasons for such independence: natural and engineered independence.[109]

Bob's orthogonal improvement is an example of a broader category of orthogonal improvements in which sets of properties are naturally independent. An earlier inventor invents and patents a new mechanical device—say, a plastic gizmo that controls Venetian blinds—that has some nonobvious interrelation among its elements. A later inventor invents a new material—say, a more durable, cheaper type of plastic. Venetian-blinds gizmos that are improved because they are made out of the new plastic are orthogonal improvements. More broadly, the literal scope of patents on advances in the mechanical arts routinely extends into mechanical devices that are improved because they are made out of after-arising materials.[110] Similarly, imagine a patent on a pill that is formed into a new shape that is easier to swallow. The literal scope of a claim to a "pill" or

---

[105] *See supra* Part I.A.

[106] *See supra* text accompanying notes [Part II.A].

[107] *See supra* text accompanying notes [Part II.A].

[108] *See supra* text accompanying notes [Part II.C].

[109] The term "natural" is not intended to invoke in any manner the products of nature doctrine or the prohibition on patenting natural principles.

[110] *See supra* note 12.

"pharmaceutical compound" molded into the specified shape would likely expand over time to encompass after-arising pharmaceutical compounds and pills made out of them. In these cases, the earlier advances are embodied in one set of properties of a thing—call them the properties of shape or geometry. The later advances are embodied in a property that can vary independently (within limits)—call it materiality.[111] Shape and materiality are to some degree naturally independent properties of a thing. One can make two things that differ in shape without requiring any difference in materiality; one can make two things that differ in materiality without requiring any difference in shape. Because the two successive advances are embodied in things as properties that can vary independently, the two advances compound with each other in the improvement.[112] The eventual embodiment of the later advance in the improvement in no way supplants or supersedes the embodiment of the earlier advance. The properties in which the advances are manifest are nonrivalrous by nature: the fact that a thing embodies one property (e.g. a new geometry) does not affect the ability of the same thing to change in the way that is required to embody the other property as well (e.g., a new materiality).

Shape and materiality are, within limits, naturally independent properties in the mechanical arts.[113] In other technologies, however, the independence of two sets of properties is not natural but engineered. This is the case in the computer industry. Hardware and software are normally developed separately. The functional properties of a programmed computer are largely independent of the physical properties of the computer that executes the program. The gates and switches can be shuffled and reshuffled, and yet the computer—whatever its final internal configuration—can still perform the same software-scripted functions.[114] The independence of the functional capacities of software and the physical characteristics of computers is not a natural phenomenon. It has been engineered through a series of technical standards and intermediary technologies such as compilers. Because software and hardware have this engineered independence, computer-related technologies give rise to many orthogonal improvements. Imagine an earlier "apparatus" claim to a software invention and the later development of new computer hardware on which the software can be run. The new hardware executing the earlier-claimed software is an orthogonal improvement. No argument of non-infringement of an apparatus claim to a software invention has ever been raised alleging that the defendant's device runs on later-developed hardware, despite the twenty-year

[111] The variation must be within limits. For example, the mousetrap itself cannot have the materiality property of *being made of cheese*.

[112] "Compound" is used here loosely in the pharmaceutical sense of the word—to mix two entities together so that they cannot be separated. [check this]

[113] They are not, however, independent properties in the chemical arts insofar as inventors claim molecules-as-things. This difference may help to explain why the temporal reach of patent scope into improvements is not as significant in the chemical arts as it is in the mechanical arts.

[114] Again, the reshuffling can only be done within limits. Many configurations of gates and switches—including anything that I could make—cannot be exchanged for a functioning computer.

term of a patent and the far less than twenty-year effective life of computer hardware.[115] The inverse scenario works as well. Imagine an earlier claim to hardware, and the later development of a new software program. It is black-letter law in the area of software patents that a machine programmed with new software is a new machine for the purposes of novelty.[116] Extending this rule to the realm of infringement, a computer programmed with the nonobvious software is an after-arising thing in relation to the disclosure of the earlier computer patent. More specifically, it is an orthogonal improvement. Again, no argument has ever been raised that earlier-patented computer hardware is beyond the temporal reach of the hardware patent just because it is a new thing in the sense that it is executing later-developed software.

2. The Proportionality Principle

One of the fundamental structuring principles of patent protection under an incentive-to-invent theory is the *proportionality principle*: inventors' patent rewards should be proportional to the social value of the ideas that inventors generate.[117] The proportionality principle mandates that the temporal reach of a patent should routinely encompass orthogonal improvements, but it is ambivalent on the temporal reach of patent scope into classic improvements.

Patents are not the only solution to the problem of inefficiently low incentives to invest in the production of technological knowledge. Government-funded research grants or government-administered prizes could also help to alleviate the problem.[118] However, one of the principal virtues of patent protection is that it decentralizes decision making about what investments should be made and turns it over to an indirect market for new ideas.[119] To possess this virtue and achieve efficiency in the allocation of investments in innovation, patents must offer stronger incentives to invest in projects that generate more valuable ideas, holding risk and cost constant.[120] The magnitude of patent incentives must be roughly proportional to the social value of the set of things embodying a newly discovered idea as measured by consumers' willingness to pay for those things. The

[115] [cites]

[116] [cites]

[117] [cites] *Cf.* Kevin Emerson Collins, *Constructive Nonvolition in Patent law and the Problem of Insufficient Thought Control*, 2007 WISC. L. REV. 759, 766–69 (discussing the "reward principle" and its proportionality limitation).

[118] [cites]

[119] The market is indirect because it does not treat ideas *per se* as property but rather treats things that embody ideas as property. *See supra* Part I.A.

[120] *See* SCOTCHMER, *supra* note [], at 97 ("Since the profit available from exclusive control of the innovation will be correlated with its social value, intellectual property encourages potential innovators to screen their ideas by comparing cost to some measure of expected social value.").

proportionality principle therefore requires patent protection to accurately track the set of things that actually embody a newly discovered idea. To achieve this goal, patent doctrine can sanction differential treatment of classic and orthogonal improvements.

The proportionality principle mandates that earlier-filed patents have sufficient temporal reach to encompass later-developed orthogonal improvements. When the later advance proceeds along an axis that is orthogonal to the axis that registers the advance that justifies the issuance of a patent, the thing embodying the later advance still by definition fully embodies the earlier advance even at its lowest level of generality.[121] Abbey's mousetrap design is still Abbey's mousetrap design even when it is made out of Bob's alloy; a software invention is still the same software invention when it is executed on after-arising hardware. The violation of the proportionality principle that would result from insufficient temporal reach into improvement is most egregious when the orthogonal improvement is a substitute for the earlier-patented thing, which is the most common scenario. In this situation, competition with the non-infringing improvement means not only that the earlier patentee would not get any profits from the improvement but also that she must lower the price on the things that she actually disclosed to the PHOSITA at the time of filing.[122]

In contrast, the proportionality principle does not provide a clear-cut answer to the "when" question of the temporal reach of patent scope into classic improvements.[123] In a classic improvement, the later advance supplants the earlier advance when the latter is framed at a lower level of generality.[124] The classic improvement therefore does not fully

---

[121] *See supra* text accompanying notes [Part III.D.1] (noting that, in an orthogonal improvement, the embodiment of the later advance does not supplant the embodiment of the advance underlying the earlier patent). Orthogonal improvements are also unlikely to be the result of independent invention, increasing the strength of the earlier patentee's normative claim to the improvement. *Cf. supra* note [] (discussing independent invention). If the after-arising advance is in an art that is different from the art the earlier invention, it is extremely unlikely that the later inventor independently invented the earlier invention. Bob, an expert in metallurgy, is unlikely to have thought up Abbey's mousetrap design on his own.

[122] In other words, when an improvement is a substitute for the things disclosed in an earlier patent, a shallow patent that doesn't encompass an orthogonal improvement doesn't simply fail to augment a patentee's profits to reflect the social value of her advance as embodied in the improvement, it eliminates the patentee's ability to profit from the technology that existed before the improvement came into existence.

It is, however, possible for an orthogonal improvement to be a complement to the things disclosed to the PHOSITA in the earlier patent specification. Imagine a world in which only drug A is known. A first inventor invents and patents pills having a new shape that is easier to swallow. A second inventor invents drug B which improves the efficacy of drug A. A pill of drug B molded into the patented shape is an orthogonal improvement, but it is a substitute for a pill molded into the patented shape made of drug A.

[123] It is in large part for this reason that the conventional wisdom on patent depth for improvements implicates so many factors. *See supra* Part II.B.

[124] *See supra* text accompanying notes [Parts II.A & D.1].

embody the advance that justified the issuance of the patent. It does embody a general advance underlying the earlier-issued patent, as the more specific after-arising advance compounds with the earlier advance at a high level of generality. [125] However, identifying the level of generality at which the embodiment of an idea in a thing should merit patent protection is a complex issue, for both normative and evidentiary reasons. As a normative matter, claims that describe inventions at high levels of generality are likely to be extremely broad and deep, meaning that they will impose more significant costs on consumers, both static and dynamic.[126] At some point, these greater costs are likely to outweigh the benefits of the additional incentives that they provide. As an evidentiary matter, it is often very difficult to determine if an inventor actually conceived of his invention at a high level of generality, or if, when described at that high level, there is no invention at all. In the mousetrap hypothetical, it was taken as a given that Abbey conceived of the very idea of the spring-loaded mousetrap.[127] However, it is possible—and, in fact, readily conceivable—that the idea of a spring-loaded mousetrap had been "in the air" for many years before Abbey finally figured out how to make one. In a possible world, Abbey could have been the first person to make a working spring-loaded mousetrap. She could have been the first person to figure out the mechanics of getting the contraption to catch mice. However, to say that she was the first person to conceive of the spring-loaded mousetrap—and to say that she deserves under the proportionality principle a slice of the profit from every improver who generates an improved spring-loaded mousetrap, regardless of how it differs from the mechanics of Abbey's mousetrap design—would require an evidentiary showing that is currently not required in order to obtain a patent.

3. Technical Advances as Economic Complements

A multidimensional model of invention-space also allows the economic concepts of substitutes and complements to be brought to bear on the decisions involved in the crafting of the temporal reach of patent scope. This represents a significant shift as, to date, these concepts have been used in patent-related doctrine only in a manner that is extrinsic to the determination of patent validity and scope. Part II.D.3.a describes the conventional wisdom on substitutes and complements in patent law. Part II.D.3.b demonstrates that, in a multidimensional invention-space, orthogonal improvements are revealed as indivisible bundles of complementary advances. To treat the successive inventors of complementary advances in the same manner in which the successive inventors of complementary things are already treated, the temporal reach of patent scope must routinely encompass orthogonal improvements.

[125] *See supra* text accompanying notes [Parts II.A & D.1].

[126] [TJ Chiang's abstraction paper?]

[127] *See supra* text accompanying notes [Part II.A].

a. Conventional Wisdom on Substitutes, Complements, and Patents

In the every-day sense of the word, "substitutes" are goods that can replace or fill in for each other. Substitute goods satisfy the same consumer need, so consumers are willing to replace one good with the other. Nails and screws are substitute goods in this common-sense way: I can use either to hang a picture, but after I opt to use one, I am less likely to use the other. In contrast, the every-day meaning of "complements" is a set of goods that consumers tend to consume together. The taste of peanut butter complements—or goes well with—the taste of jelly, so peanut butter and jelly are complements. Similarly, bitter coffee and sugar are complements: the more bitter coffee I drink, the more sugar I consume. To determine whether goods are complements or substitutes as a technical matter, economists measure the goods' cross-elasticity of demand.[128] Two goods are substitutes if a decrease in the price of one good results in a decrease in demand of the other good and, inversely, an increase in the price of one good results in an increase in demand for the other good.[129] Because consumers are willing to use either one good or the other to fulfill their needs, a decrease in the price of one will drive consumers toward that good and away from the other. The cheaper nails are, the less likely I am to purchase a screw. In contrast, two goods are complements if a decrease in the price of one good results in an increase in the demand for the other good and, inversely, an increase in the price of one good results in a decrease in the demand for the other good.[130] Because consumers tend to consume complementary goods together, the price that drives consumer purchasing decisions is the price of the bundle of goods. A decrease in the price of one good in the bundle decreases the price of the bundle as a whole, meaning a consumer will tend to consume more of the bundle and thus more of the other good. The cheaper coffee is, the more sugar I will consume (presuming the price of sugar is stable).[131]

The economic concepts of complements and substitutes are not foreign to patent scholarship.[132] For example, the concepts of substitute and complementary things play an important role in understanding patent incentives and rewards in cumulative

---

[128] [define elasticity & cross-elasticity]

[129] VARIAN, *supra* note [], at 111.

[130] *Id.* at 112.

[131] Perfect substitutes exist "if the consumer is willing to substitute one good for the other at a *constant* rate." *Id.* at 38. Perfect complements are "goods that are always consumed together in fixed proportions." *Id.* at 40.

[132] The notion of complementary is not the one used by Bessen & Maskin in their article on cumulative innovation. Bessen & Maskin, *supra* note [], at 612 ("[B]y 'complementary,' we mean that each potential innovator takes a different research line and thereby enhances the overall probability that a particular goal is reached within a given time.").

innovation.[133] If an earlier-invented thing and a later-invented thing are neither substitutes nor complements, then the incentives and rewards provided by the patent regime for each invention are entirely independent. If they are substitutes, then the invention of the later thing may radically decrease the patent profits of an earlier inventor. Consider the common notion of "design around" in which a later inventor successfully produces a thing that is a substitute for an earlier patentee's technology and that lies beyond the scope of the earlier patentee's claims.[134] Design-around has a significant negative impact the ability of a patentee to charge supra-competitive prices for the technology that she actually enabled, demonstrated possession of, and disclosed in her specification. However, if the earlier-invented thing and the later-developed thing are complements, the earlier patentee's profits will, if anything, increase at the time of the later invention. Consider the situation when an earlier inventor has invented the hammer and a later inventor develops an improved nail.[135] The development of an improved nail that is easier to pound without being bent increases the value of existing hammers to consumers, creating a surplus that must be split in some manner between consumers, the hammer-inventor and the improved-nail-inventor. Precisely how much of the potential surplus is realized and precisely how the realized surplus will be split is a complicated issue that turns on both the patent landscape and the efficiency of the market for welfare-enhancing licensing schemes.[136] Precisely how the increase in the value of an earlier-patented good (whether it is a thing or an embodied advance) upon the later discovery of a complementary good should be distributed between the earlier and later inventions is also complicated issue.[137] Under no circumstances, however, can the owner of the hammer patent be made worse off by the later development of the improved nail.

The concepts of substitutes and complements also play an important doctrinal role in the scrutiny of patent licenses for anticompetitive behavior, but here it is patent claims that are identified as either substitutes or complements. Patent claims are labeled as substitutes or complements in either of two ways. First, if there is no overlap in the scope

---

[133] *Cf*, Michael Abramowicz, *Perfecting Patent Prizes*, 56 VAND. L. REV. 115, 155–57 (2003) (discussing how substitute and complementary products distort patent incentives).

[134] *See supra* text accompanying note [] (defining "design around"). The patent regime creates a strong incentive for firms to design around their competitor's patents. Abramowicz, *surpa* note [], at 156. Courts frequently laud these incentives to design around as one of the virtues of the patent regime, but economic-minded commentators have understandably mixed opinions on the subject. []

[135] Another common example occurs in the computer industry: later-developed software applications make earlier-patented hardware devices and operating systems more valuable.

[136] If the owners of patents on complementary goods make independent pricing determinations, then there is inefficient over-pricing of the bundle even from the perspective of the profit-maximizing monopolists. Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting*, 1 INNOVATION POLICY & THE ECONOMY 122–24 (2000).

[137] The optimal division of this surplus is a topic that is beyond the scope of this Article.

of the two claims, then the status of the claims as complements or substitutes simply piggybacks on the status of the things described by the patent claims. Just as nails and screws are substitutes, patents on nails and screws are substitutes. Just as hammers and nails are complements, a patent on a hammer and a patent on a nail are viewed complements.[138] Second, if the claims do overlap, then the patents are labeled as complements. Both patent rights are required to make, use, or sell the self-same thing. Because any consumer who desires to use that thing must acquire rights to practice both patents, the patents are by definition consumed together and are therefore complements. The distinction between substitute and complementary patents is important because it helps to identify patent licenses that are in effect agreements not to compete.[139] If patents are substitutes, then an exclusive agreement to cross license the patents can be agreements not to compete on price. However, if patents are substitutes, patent licenses are likely procompetitive.

b. Orthogonal Improvements Are Indivisible Bundles of Complementary Advances

The conventional wisdom on substitutes and complements identifies things (and derivatively patents) as the relevant goods. A shift away from things is required in order to use the concepts of substitutes and complements to craft the temporal reach of patent scope. Advances, or more precisely the properties that embody advances in things, can be either complements or substitutes, as well. As with things, substitute and complementary advances are defined by cross-elasticity of demand. Two advances are substitutes if an increase in the price of the right to "use" the property that embodies one advance leads to an increase in demand for the right to "use" the property that embodies the other advance.[140] For example, the hypothetical advances of square and round pills

[138] Sometimes, things covered by claims are both complements and substitutes. For example, imagine two patents on two different shapes for pills. Whether pills-as-things are complements or substitutes is a fact that turns on the chemical composition of the pill, not the shape. Therefore, pills that are both complements and substitutes will fall within both patents. In this situation, the nature of the patents as complements or substitutes mirrors the nature the hypothetical rights to properties that are required to identify advances as complements or substitutes. *See infra* note 140.

[139] 2 HERBERT HOVENKAMP, MARK D. JANIS, & MARK A. LEMLEY, IP AND ANTITRUST § 34.2c, at 34-6 to -10 (2004); DEPT. OF JUST. & FTC, ANTITRUST GUIDELINES FOR THE LICENSING OF INTELLECTUAL PROPERTY § 5.5 (1995), *available at* www.usdoj.gov/atr/public/guidelines/0558.htm (discussing the anticompetitive and procompetitive effects of cross-licensing and patent pooling arrangements); Shapiro, *supra* note [], at 119.

[140] To identify embodied advances as complements or substitutes, it is necessary to imagine a hypothetical property regime in which manufacturers must pay for the right to embody an advance in a thing and/or technology users must pay for the right to use things embodying the advance. If embodied advances are not treated as hypothetical commodities for which consumers must pay, then the very idea of cross elasticity of demand is nonsensical. If the advance is the idea of making circular pills, then anyone who makes, uses, or sells pills that possess the property of being circular consumes the embodied advance

These hypothetical property rights resemble patent rights, but, importantly, they are not necessarily the rights that would actually be granted by the contemporary patent regime upon the discovery of the advance.

are substitutes because the demand for square pills will increase if it costs more to manufacture round pills. Similarly, two advances are complements if an increase in the price of the property that embodies one leads to a decrease in demand for the property that embodies the other. There are two distinct reasons why two advances may be complements that parallel the two distinct reasons why patents are identified as complements in the scrutiny of patent licensing for anticompetitive behavior.[141] On the one hand, the two properties define the characteristics two distinct things that make consumers treat those things as complements. If chemicals A and B are complements because of the properties with which their chemical structures endow them, then the property *having the molecular structure of chemical A* and the property *having the molecular structure of chemical B* are complements. On the other hand, the two properties may be complements if a single, indivisible thing can embody both properties and consumers desire that thing. Consider a pill that is molded into a shape that is easy to swallow and that is made out of chemical X. A single, indivisible thing can embody both properties. Furthermore, consumers desire the bundle of the properties that the thing embodies. Consumption decisions will be based on the price of the bundle. The price of the bundle is determined by the sum of the prices of the individual embodied advances. Therefore, an increase in the price of one good will lead to a decrease in the demand for the other as demand for the bundle decreases.

It is this second situation which is relevant to orthogonal improvements. The successive advances embodied in an orthogonal improvement are by definition economic complements. Demand for the improvement-as-thing is desire for the bundle of the earlier and later advances that are both fully embodied in the orthogonal improvement. If the price of the right to "use" the property embodying one advance goes up, demand for the right to "use" the property embodying the other advance will go down. The bundle of the properties in a thing just happens to be indivisible in that it is impossible to separate out a property as an ontological entity and sell it as an entity distinct from a thing that possesses the property. It is, however, still a bundle.

Once the successive advances in an orthogonal improvement are identified as complements, there is a straight-forward justification for allowing the temporal reach of patent scope to routinely encompass orthogonal improvements. There is no reason for the patent regime to treat successive inventors of complementary advances that are embodied in things as properties any differently from the way the patent regime treats successive inventors of complementary things.

The discovery of an after-arising complementary thing (an improved nail) increases the social value of an existing thing (an existing hammer). Similarly, the after-arising advance that allows springs to be made out a cheaper, stronger alloy increases the social value of the earlier-patented advance in the arrangement of the mechanical

---

Because they identify only a property of a thing, rather than a set of things, they can reach infinitely far into future technologies. Such depth in the contemporary patent regime cannot be taken for granted.

[141] *See supra* text accompanying notes [Part II.D.3.a].

components of a spring-loaded mousetrap. To the same end, the later-developed property of a pill *being made out of chemical X* increases the social value of the property *being shaped such that being swallowed is easier*.

When an after-arising thing is a complement to an earlier-patented thing, the earlier patentee can profit from the after-arising complement even if the earlier patent has no temporal reach at all. A patent on A reads on the bundle of A+B because the A in A+B is the same old A that was disclosed in the A patent. Yet, for the private value of Abbey's patent will mirror the private value of the hammer patent upon the discovery of an after-arising complement, Abbey's patent must have sufficient temporal reach to encompass orthogonal improvements. In the ontology of things, an orthogonal improvement that indivisibly bundles together complementary properties is an A', not an A+B.[142] Abbey's patent will therefore encompass Bob's mousetrap only if the set of things that it encompasses can expand beyond the set of things disclosed to the PHOSITA at the time of filing and encompass after-arising improvements. Where the hammer and nail patentees can obtain distinct-thing blocking patents, Abbey and Bob can get blocking patents only if there is sufficient depth to Abbey's claim to produce same-thing blocking patents.[143] If earlier patents do not have sufficient temporal reach to encompass orthogonal improvements, then an event that in theory increases the social value of an earlier-patented invention will radically decrease the private rewards that the patent regime offers to the earlier inventor.[144] Failing to allow earlier patents to encompass orthogonal improvements would be economically akin to saying that it is not an infringement of an earlier patent on a hammer to make, use, or sell hammers that are employed only to pound on after-arising nails.

E. Coda: Revisiting Merges and Nelson

Published twenty years ago, *On the Complex Economics of Patent Scope* by Robert Merges and Richard Nelson was a groundbreaking article on the how patent protection should be tailored to most effectively foster cumulative innovation.[145] It set the terms of the debate for much of the debate over patent protection for cumulative innovation during the last two decades.[146] With so much commentary on the Merges and Nelson article now layered on top of it, the article itself is worth revisiting and examining

[142] *See supra* [Parts I.B & C].

[143] *See supra* text accompanying notes [Part I.C] (distinguishing same-thing and distinct-thing blocking patents).

[144] This result presumes that the orthogonal improvement is a substitute for the things disclosed in the earlier patent. *But see supra* note [] (noting that orthogonal improvements can be complements).

[145] 90 COLUM. L. REV. 839 (1990).

[146] [string cite]

in a new light. More specifically, although the existence of orthogonal improvements has long been overlooked, the distinction between classic and orthogonal improvements developed in this Article is already implicit in, or at least strongly foreshadowed by, Merges and Nelson's distinction between cumulative and discrete technical advance.

Merges and Nelson articulate a thesis about the correlation between optimal patent scope and the pattern of innovation in an industry as either discrete or cumulative.[147] "Discrete" innovation industries are industries in which an invention "does not point the way to wide ranging subsequent technical advances."[148] In contrast, "cumulative" innovation industries are industries in which "today's advances build on and interact with many other features of existing technology."[149] Merges and Nelson are concerned about the impact of broad, early patents on the pace of later improvements in cumulative innovation industries but not in discrete innovation industries. "Particularly when the technology is in its early stages, the grant of a broad-gauged pioneer patent to one party" in a cumulative innovation industry "may preclude other inventors from making use of the inventions without infringing the original patent."[150] In contrast, Merges and Nelson "are not interested in the cases of discrete invention" because they "are concerned with the effects of patent scope decisions on the subsequent development of technology" and presumptively, when patents describe discrete technologies, there are no subsequent developments into which to reach and that need to be open to other inventors to make.[151] Merges and Nelson therefore discuss cumulative technologies at length over thirteen pages,[152] but they mention discrete innovation only in passing to identify it and put it aside.[153]

There are aspects of the distinction between cumulative and discrete technical advance made by Merges and Nelson that the framework of this Article singles suggests are in need of further refinement. For example, in identifying cumulative innovation industries, Merges and Nelson make no distinction between industries that are cumulative because (a) components developed by earlier inventors are used by later inventors in

---

[147] *See supra* Part II.B.2. Merges and Nelson also separate out chemical and science-based industries for distinct treatment. [] The categories are derived from Nelson's earlier work. []

[148] Merges & Nelson, *supra* note [], at 880.

[149] *Id.* at 881.

[150] *Id.* at 882. The reasoning underlying this normative conclusion is explained in the text accompanying *supra* notes [Part II.B.2].

[151] Merges & Nelson, *supra* note [], at 884.

[152] *Id.* at 884–97.

[153] *Id.* at 884.

conjunction with complementary things in larger assemblies and (b) pioneering patents describe early-stage discoveries that can be developed in many different directions.[154] It is true that patents granted to earlier inventors in industries with either of these characteristics pose difficulties for later inventors, and that the economics of patent protection needs to recognize the impact of patent protection on cumulative innovation in both contexts. However, if the goal is to identify how patent scope can be tailored to facilitate cumulative innovation,[155] then lumping these two problems together is problematic. The complementary components issue cannot be addressed through the tailoring of the temporal reach of patent scope, as even very narrow, shallow patents—the very narrow, shallow patents needed to prevent outright piracy of the things that a patent discloses to the public—will trigger inter-generational patent fights. Even the narrowest of peripheral patents on A must read on A+B in an effective patent regime, regardless of whether B is after-arising technology.[156] In contrast, the issue of early-stage, pioneering developments can be effectively addressed through doctrinal proposals aimed at curtailing the reach of patent scope into improvements along both the dimensions of breadth and depth.[157]

There are other aspects of the distinction between cumulative and discrete technical advances that Merges and Nelson took care to note but that subsequent scholarship has since glossed over. Merges and Nelson were clear that different patterns of advance may predominate in different industries but that a pure industry-by-industry classification is an overgeneralization.[158] At its bottom, the question about tailoring of patent scope is an invention-by-invention, patent-by-patent, and improvement-by-improvement decision in which industrial classifications are only rough proxies. Additionally, Merges and Nelson were careful to note that even discrete technical advance could involve improvements, patents with depth, and blocking patents.[159]

---

[154] *Id.* at 881–82.

[155] *Id.* at 916 ("Our goal has been to show that scope doctrines can be used to approximate the 'tailoring' function proposed by economists who model optimal patent length, with an eye toward retaining incentives for subsequent improvements.").

[156] *See supra* [Part I.B & C] (distinguishing component re-use from improvement). The primary policy lever for dealing with inter-generational component problems may be remedies, not claim scope. [cite on royalty stacking & injunctive relief]

[157] In identifying the doctrines that can restrict claim scope, Merges and Nelson work with enablement, the doctrine of equivalents, and the reverse doctrine of equivalents. *Id.* at 845–68. In the terminology used in this Article, enablement can constrain claim depth, *see infra* [Part III. B.1], whereas the doctrine of equivalents and the reverse doctrine of equivalents constrain claim breadth, *see infra* [Part III.B.4]. This Article also enriches the set of policy tools that can be used to craft claim depth. *See infra* [Part III.C].

[158] Merges & Nelson, *supra* note [], at 880 ("In any industry one or another of these models may be applicable at any given time, or appropriate characterization may require a mix.").

[159] *Id.* at 880, 882.

However, perhaps because they did not devote a section to discussing discrete advance technologies, or perhaps because the name "discrete" is subject to misreading, one between-the-lines interpretation of Merges and Nelson has been that after-arising improvements simply do not occur when discrete technologies are at issue.[160] This is clearly a misreading of Merges and Nelson's thesis. The absence of any extended discussion about patent scope for discrete technical advance in Merges and Nelson is not because patents on discrete inventions never encompass after-arising technology but rather because whatever temporal reach these patents achieve over their lifespan is not problematic as a normative matter.

Putting all of these observations together, the distinction between classic and orthogonal improvements articulated in this Article is arguably simply a reworked version of Merges and Nelson's distinction between cumulative and discrete technologies. Recognizing that the components problem raises a set of problems for patent protection that are unrelated to patent scope *per se*, the classic/orthogonal distinction carves the problem off and places it to one side.[161] Rather than looking outward to industrial context and working on an industry-by-industry basis, the classic/orthogonal distinction looks inward to the facts of specific cases and attempts to identify two different scenarios in which later improvers improve on the things claimed by earlier inventors. Like Merges and Nelson's distinction, the classic/orthogonal distinction identifies one type of improvement that earlier patentees can control without raising any red flags (orthogonal) and one type that would be problematic if excessively controlled by earlier patentees (classic). Although the thesis lies beyond the scope of this Article to prove, there is likely a strong correlation between patents likely to develop orthogonal improvements and patents covering discrete advance technologies. Inversely, patents that are highly susceptible to classic improvements may prove to be patents on cumulative advance technologies.

There is more than one way to skin a (thankfully conceptual) cat, and the classic/orthogonal distinction may simply be another way of getting at the cumulative/discrete distinction set forward by Merges and Nelson twenty years ago. Rather than identifying which types of patents are least and most likely to raise normative problems when construed so as to have considerable reach into improvements, however, it identifies which types of improvements are least and most likely to raise normative problems when routinely encompassed within claim scope. This latter perspective is critical if answers to the "when" question of temporal reach are to be bridged over into the doctrinal "how" question as Part III attempts to do.

---

[160] Christopher A. Cotropia, *"After-Arising" Technologies and Tailoring Patent Scope*, 61 N.Y.U. ANN. SURV. AM. L. 151, 154 (2005) (arguing that firms in discrete innovation industries "have little worry about being substituted by follow-on inventions").

[161] *See supra* text accompanying notes [Parts I.B & C].

III. RETHINKING THE "HOW" QUESTION

The distinction between classic and orthogonal improvements also has implications for the "how" question of the temporal reach of patent scope into improvement. Part III.A presents the common *fixation theory* about the nature of peripheral claims: peripheral claims cannot have depth because they fix the set of things that fall within the literal scope of a claim to things that are disclosed to the PHOSITA on the date of filing. In order to accommodate the previously overlooked fact that that the literal scope of a contemporary peripheral claim routinely has depth in that it encompasses orthogonal improvements, Part III.B revises this theory of the peripheral claim. It draws on the philosophy of language to demonstrate that the fixation of the meaning of claim language to the PHOSITA on a historical date does not entail fixation of the set of things described by the claim language to the set of things known to the PHOSITA on that date. In brief, the fixation of *ideational* meaning entails freezing word-to-word relationships, not word-to-world relationships. Part III.C considers the implications this revised, ideational theory of the peripheral claim, when combined with the observation in Part II that patent depth should routinely encompass orthogonal improvements but not classic improvements, for four patent doctrines that courts use to curtail both the patent breadth and depth for improvements. Part III.D raises a final implication of the revised theory: patent claims are both peripheral and highly dependent on the spirit or essence of the patentee's invention.

A. The Fixation Theory of Peripheral Claims

The contemporary patent regime is a *peripheral* claiming regime in which the patent applicant provides a textual description of the set of claimed things and the "outer boundary" of a patent claim is therefore fixed as of the date on which the patent is filed.[162] In contrast, a *central* claiming regime would simply publicize a prototypical example of an invention and determine the extent of a claim's scope on a case-by-case

[162] *Supra* note 30. Jeanne Fromer suggests a distinction between peripheral claiming "by exemplar" in which every member of the claimed set is presented individually and peripheral claiming "by characteristic" in which the properties of things that are common to all members are described. Fromer, *supra* note [], at 727. This Article addresses only peripheral claims in which the outer boundary is delineated by a descriptive text—a subset of what Fromer refers to as peripheral claiming by characteristic. Whether true peripheral claiming by exemplar is a viable claiming technique is highly questionable. Fromer recognizes the problem that occurs when membership in the claimed set exceeds a very small number. *Id.* at 760. However, even small-sized peripheral claims by exemplar are problematic. Consider the best-case scenario for a peripheral claim by exemplar: a claim to a newly fabricated molecule in which a three-ounce sample is given to a PTO-approved depository in a vial. In what ways do other substances need to be just like the exemplar? Geographically? Contained in a duplicate vial? In sample-sizes of three ounces? *See generally* [Wittgenstein, rabbit parts ditty]. It is always necessary to identify the characteristics of an exemplar-thing that must be present in the accused device for there to be infringement.

basis by querying whether an allegedly infringing technology is sufficiently similar to the exemplar to merit infringement.[163]

One commonly assumed attribute of a peripheral claim is that the set of things described by a peripheral claim is rigidly fixed to the set of things disclosed to the PHOSITA on the date the claim is filed. The negative corollary of this assumption is that a peripheral claim cannot encompass after-arising technology such as an improvement.[164] Together, the assumption and its negative corollary yield the *fixation theory* of peripheral claims.

The fixation theory is most commonly defended as a normative matter with two arguments. The first is that peripheral claims that adhere to the fixation theory generate more effective public notice of claim scope. While there are costs and benefits of both peripheral and central claims,[165] the primary justification of peripheral claims in general lies in the greater certainty and notice to the public that they generate concerning the scope of the set of claimed things.[166] The fixation theory picks up this ball and runs with it, arguing that fixation of the full set of things within the scope of a claim on the date of filing is required for the public to have effective notice of the scope of the claim.[167] The second justification of the fixation theory is an interpretation of the proportionality principle:[168] if patentee's rights are to be proportional to what they invented, patentees'

---

[163] For descriptions of the difference between central and peripheral claiming, see generally Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts? Rethinking Patent Claim Construction*, 157 PENN. L. REV. 1743 (2009); Fromer, *supra* note [].

[164] *See infra* notes [] and accompanying text (discussing the doctrinal and normative support for the fixation theory). One manifestation of the fixation theory is an economic model of patent protection that presumes patents have no depth because it equates the novelty or nonobviousness of a subsequent invention with the noninfringement of an earlier patent. Robert M. Hunt, *Patentability, Industry Structure, and Innovation*, 12 J. INDUS. ECON. 401, 404 (2004) (modeling patent rights in which "the inventor may use all prior discoveries without licensing them" if the inventor has produced "an invention [that] satisfies the standard of" nonobviousness); Suzanne Scotchmer and Jerry Green, *Novelty and Disclosure in Patent Law*, 21 RAND J. ECON. 131 (1990). [Durham on patent symmetry]

[165] Burk & Lemley, *Sign Posts*, *supra* note [], at []; Fromer, *supra* note [], at [].

[166] Burk & Lemley, *Sign Posts*, *supra* note [], at 1745 n.11. [same as generic argument for rule over standard]

[167] *See* JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 67 (2008) (linking a lack of clarity in the meaning of claim terms to the ability of claims to encompass after-arising technology).

[168] *See supra* [Part II.D.2] (discussing the proportionality principle).

claims should not encompass things that the patentee did not invent or disclose to the PHOSITA at the time of filing.[169]

The fixation theory finds direct support in two patent doctrines. In part, it is grounded in the Federal Circuit's case law on claim construction. To avoid the uncertainty that would result from allowing the meaning of claim language to change over time, the legal meaning of claim language is anchored to the meaning that was understood by the PHOSITA at the time of filing.[170] The Federal Circuit has at times stated that the fixation of the meaning of claim language on the date of filing entails the fixation of the set of claimed things to the set disclosed to the PHOSITA on the date the patent is filed.[171] The implication is that literal claims therefore "cannot capture the later-developed technologies because to do so would require the claims to be interpreted as they are understood at some time after the filing date."[172] In part, the fixation theory is also grounded in judicial pronouncements on the disclosure doctrines of enablement and written description. Enablement requires the set of claimed things to be commensurate with the set of things that the disclosure teaches the person having ordinary skill in the art (the PHOSITA) how to make and use at the time of filing.[173] If construed strictly under a

[169] *See* O'Reilly v. Morse, 56 U.S. (15 How.) 62, 113 (1853) (invalidating a patent claim because it was "too broad" because the inventor "claims an exclusive right to use a manner and process which he has not described and indeed had not invented, and therefore could not describe when he obtained the patent"); Superguide Corp. v. DirecTV Enters. Inc., 358 F.3d 870, 898 (Fed. Cir. 2004) (Michel, J., concurring) (""[T]he applicant must be the 'inventor' of the things covered by the patent claims."); Schering Corp. v. Amgen Inc., 222 F.3d 1347, 1353–54 (Fed. Cir. 2000) (""Because, at the time of the . . . application . . . [those] skilled in the art [did not know] of the existence of [the allegedly infringing AAT, the AAT] cannot be within the scope of the claims . . . . To grant broader coverage would reward [the inventor] for inventions he did not make."); Holbrook, *supra* note [], at 7 (arguing that there is a "paradox" involved in "affording patent protection for devices the inventor did not create"); O'Donoghue, Scotchmer, & Thisse, *supra* note [], at 3 (noting that claim depth "seems prima facie untenable on the basis that it gives property rights on qualities (products) that the patent holder did not invent").

[170] Phillips v. AWH Corp., 415 F.3d 1303, 1313 (Fed. Cir. 2005) (en banc).

[171] Chiron Corp. v. Genentech, Inc., 363 F.3d 1247, 1263 (Fed. Cir. 2004) (Bryson, J., concurring) ("I think the proper approach . . . is to address cases of new technology by construing claims, where possible, as they would have been understood by one of skill in the art at the time of the invention, and not construing them to reach the as-yet-undeveloped technology . . . ."); Superguide Corp. v. DirecTV Enters., 358 F.3d 870, 896 (Fed. Cir. 2004) (Michel, J., concurring); Schering Corp. v. Amgen Inc., 222 F.3d 1347, 1349–50, 1353–54 (Fed. Cir. 2000) ("Because, at the time of . . . application, neither [the inventor] nor others skilled in the art knew of the existence of, let alone the identity of, the specific polypeptides now identified as subtypes of IFN-α, those subtypes cannot be within the scope of the claims.").

[172] Cotropia, *supra* note [], at 165; *id.* at 167–68 ("[R]ecent case law has emphasized the temporal limitation on literal claim meaning, prohibiting the literal capture of later-developed technologies."). *See also* Mark A. Lemley, *The Changing Meaning of Patent Claim Terms*, 104 MICH. L. REV. 101, 116 (2005) (arguing in favor of fixing a claim's meaning on the date of filing because delaying fixation until the date of infringement "would require the scope of patents to change over time" and would mean that "a patent's scope would not be fixed, but could differ from infringer to infringer as time passes").

[173] Genentech, Inc. v. Novo Nordisk A/S, 108 F.3d 1361, 1365 (Fed. Cir. 1997).

"full scope" rule, the enablement doctrine mandates the fixation theory because the set of things encompassed within a claim cannot be larger than a set of things that the PHOSITA knew how to make and use at the time the claim is filed.[174] The written description requirement requires the set of claimed things to be commensurate with the set of things that the PHOSITA, having read the specification, recognizes as "invented" or "possessed" by the inventor at the time of filing.[175] Strict interpretations of the written description requirement, too, support the fixation theory. If the set of things encompassed within a claim is limited to the set of things that were "invented" and "possessed" by the patentee on the date a patent is filed, after-arising technologies are categorically excluded from peripheral claims.[176]

The fixation theory is also reinforced by the real-property metaphor that is commonly invoked to explain the nature of a peripheral claim. Peripheral claims, it is said, are like the "metes and bounds" of a parcel of land.[177] In the same vein, peripheral claims are described as "fences" marking a patentee's interest.[178] The land that is subject to property rights is a physical entity that occupies the identified coordinates in space, so delineating the boundary of a parcel of land entails identifying the full scope of the propertized resource *ex ante*, just as the fixation theory of peripheral claims suggests should occur in patent law.[179]

---

[174] *See* Sitrick v. Dreamworks, LLC., 516 F.3d 993 (Fed. Cir. 2008); Pharm. Res., Inc. v. Roxane Labs., Inc., 253 F. App'x 26 (Fed. Cir. 2007); Auto. Techs. Int'l, Inc. v. BMW of N. Am., Inc., 501 F.3d 1274 (Fed. Cir. 2007); Liebel-Flarsheim Co. v. Medrad, Inc., 481 F.3d 1371 (Fed. Cir. 2007); AK Steel Corp. v. Sollac, 344 F.3d 1234, 1241 (Fed. Cir. 2003). *See generally* Bernard H. Chao, *Rethinking Enablement in the Predictable Arts*, 2009 STAN. TECH. L. REV. 3, ¶¶ 23–49 (discussing recent cases employing the full-scope variant of the commensurability requirement). A related line of reasoning suggests that an enabled claim cannot encompass an embodiment that is nonobvious in relation to the disclosure. Charles W. Adams, *Allocating Patent Rights Between Earlier and Later Inventions*, 54 ST. LOUIS U. L.J. 55 (2009); [BU symmetry article].

[175] Univ. of Cal. v. Eli Lilly & Co., 119 F.3d 1559 (Fed. Cir. 1997).

[176] *See* Ariad Pharms., Inc. v. Eli Lilly & Co., 598 F.3d 1336, 1365–66 (Fed. Cir. 2010) (*en banc*) (Rader, J., dissenting) (arguing that the written description requirements makes blocking patents impossible); Chiron Corp. v. Genentech, Inc., 363 F.3d 1247, 1255 (Fed. Cir. 2004) (invalidating a claim under the written description doctrine "[b]ecause [the allegedly infringing AAT] technology did not even exist at the time of the . . . filing, the record conclusively supports that the Chiron scientists did not possess and disclose this technology").

[177] FABER, *supra* note [], at §1:2 ("A customary analogy is that the patentee must stake out the precise boundaries of his claim, in a manner similar to the description of the boundaries of a plot of land.").

[178] Burk & Lemley, *supra* note [], at 1744; Craig Allen Nard, *Certainty, Fence Building, and the Useful Arts*, 74 IND. L.J. 759, 759 (1999).

[179] Jeffrey A. Lefstin, *The Formal Structure of Patent Law and the Limits of Enablement*, 23 BERKELEY TECH. L.J. 1141, 1780 (2008).

B. A Revised, Ideational Theory of Peripheral Claims

For advocates of the fixation theory, orthogonal improvements are an inconvenient truth. Earlier patentee's interests should routinely encompass them.[180] Furthermore, earlier patentee's interests do routinely encompass them, and it is literal claim scope that most frequently does the needed work.[181] Orthogonal improvements are empirical evidence that fatally compromises the fixation theory of peripheral claims.[182]

The fall of the fixation theory, however, leaves a conceptual void for the doctrine of claim construction if the meaning of claim terms is to be fixed on the date of filing.[183] A revised, ideational theory of the peripheral claim can fill this void and explain how the outer boundaries of a peripheral claim can be stabilized on the date a claim is filed in a manner that generates effective public notice without limiting the set of claimed things to technological things known to the PHOSITA at that time.[184]

There is an ambiguity in the meaning of "meaning." On the one hand, words can have *denotational* meaning. Here, meaning is based on reference or word-to-world relationships. Meaning is a set-theoretical construct: the set of possible things in the world to which words refer *are* the words' meanings. For example, the word "fastener" means in a denotational way the set of possible objects paperclips, nails, shoelaces and glue dots to which it refers. Because denotational meaning elides meaning-scope and thing-scope, the inclusion of after-arising technology within the literal scope of a denotationally construed necessitates a change in the meaning of claim language. If inventor X files a claim that recites "fastener" as a limitation and Velcro® is invented after the filing date, the denotational meaning of inventor X's claim cannot encompass Velcro® without a change in meaning.

On the other hand, words can also have ideational meaning. Here, words acquire meaning by their existence as a node that is linked into a network of other words and

---

[180] *See supra* Parts II.D.2 & 3.

[181] *See supra* text accompanying notes [Part II.D.1]. Furthermore, it makes no sense to shift the status quo and allow orthogonal improvements to infringe only under the DOE. *See infra* text accompanying notes [Part III.C.4].

[182] Orthogonal improvements are the easiest improvements to use to make the case that the fixation theory is untenable. However, once the point is made, the revised, ideational theory of peripheral claims explains both how orthogonal improvements can routinely wind up within literal claim scope and how classic improvements can sometimes do so. *See infra* [Parts III.C.1–3].

[183] The fall of the fixation theory also challenges the "full scope" variant of the enablement doctrine and its parallel in written description, but there are other theories waiting in the wings. *See infra* Part III.C.2.

[184] The following paragraphs are summaries of the arguments made in Collins, *Things and Meanings*, *supra* note [], at 536–53 (demonstrating the import of using either ideational or denotational meaning in claim construction).

concepts that exists in our minds. Continuing the example, the ideational meaning of the word "fastener" is to be found in the link between the concept of *fastener* and *join* that specifies the relationship between the two as the former *having the function of* the latter, as well as in many other links with many other concepts. In contrast, the meaning of denotationally construed claim language, the meaning of ideationally construed claim language can remain stable even as the language refers to an ever-expanding set of objects. The mental network concept-nodes and relationship-linkages can be steadfastly fixed on the date a patent claim is filed even as new technologies described by a word are introduced into the storehouse of human knowledge. The thing-scope of an ideational claim to a "fastener" can thus expand along a progressive dimension over time and encompass the after-arising Velcro®, even as its meaning-scope remains resolutely fixed on the date of filing.

Assuming ideational meaning governs claim construction (as it usually does[185]), there is no inherent conflict the inclusion of unpossessed, after-arising technology within literal claim scope and the stable, notice-enhancing fixation of claim meaning on the date on which a patent is filed. The meanings of phrases like "carriage body," "normally flat aeroplane," "device," "apparatus," and "pill" can fixed on the date a patent is filed, and yet the patents can have depth in that they encompass later-developed improvements.

Furthermore, the ability of a patentee's claim to reach into technological things that the inventor did not invent is not problematic under the proportionality principle.[186] Patent incentives should be proportional not to the set of things created by the inventor but rather to the inventive ideas produced by the inventor that are embodied in things.[187] If patent rewards are to in any way track the social value of ideas, the set of things that fall within claim scope must be allowed to grow over time.

C. Implications for the Doctrine that Controls the Temporal Reach of Patent Scope

A revised, ideational theory of the peripheral claim enables claims with depth as a technical matter,[188] but it does not give any normative guidance on when claims should have significant depth. This guidance comes instead from the distinction between classic and orthogonal improvements articulated and examined in Part II: claims should routinely have sufficient depth to encompass orthogonal improvements, but claim depth for classic infringements should be scrutinized more closely.[189]

---

[185] *See id.* at 548–53.

[186] One justification of the fixation theory suggests that it is problematic. *See supra* text accompanying notes [Part III.A].

[187] *See supra* Part I.A.

[188] *See supra* Part III.B.

[189] *See supra* Part II.D.

This Part explores the doctrinal implications of abandoning the fixation theory of peripheral claims and using both claim depth and breadth to afford different levels of protection for classic and orthogonal improvements. In particular, it examines four patent doctrines that control patent depth and breadth: the doctrine of claim construction as articulated in *Phillips v. AWH Corp.*,[190] the disclosure doctrines, the rules of means-plus-function claiming, and the doctrine of equivalents. Importantly, however, many of the implications require a change in the status quo only in how courts describe what they are already doing, as the contemporary patent regime does not—and historically never has—adhered to the fixation theory, nor has it been insensitive to the distinction between classic and orthogonal improvements.

1. Claim Construction

A revised, ideational theory of the peripheral claim is already a theory about claim construction. Courts should, by default, fix the ideational meaning of claim terms rather than the denotational meaning of claim terms at the time of filing.[191]

However, there is another implication of the revised, ideational theory of the peripheral claim when it is taken in conjunction with the distinction between classic and orthogonal improvements. It is difficult to preserve the façade maintained by the Federal Circuit that claim construction is a policy-neutral doctrine.[192] The reach of patent scope into after-arising improvements is not—and should not be—uniform for all types of improvements. Rather, the temporal reach of patent scope should be sensitive to the type of improvement that is at issue: it should reach more readily into orthogonal improvements than it does into classic improvements. If the improvement is an orthogonal improvement, sufficient temporal reach to encompass the improvement should be routine. However, if the improvement is a classic improvement, then sufficient temporal reach to encompass the improvement should not be routine but should instead be determined on a case-by-case basis. To achieve this selectivity, however, courts must perform claim construction with one eye on the nature of the allegedly infringing improvement and work backward from the desired end result.[193]

There are, in theory, two doctrinal policy levers for modulating claim construction to achieve this goal of selectivity—through the tailoring of either claim breadth or claim depth.

---

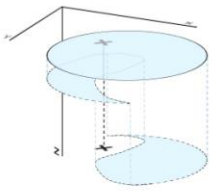
[190] Phillips v. AWH Corp., 415 F.3d 1303 (Fed. Cir. 2005) (en banc).

[191] *See supra* Part III.B.

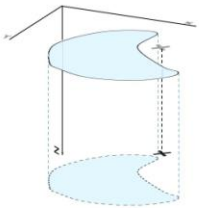
[192] The Federal Circuit commonly describes claim construction as a process that should not be affected by whether a claim should be broad or narrow as a matter of policy. []

[193] Working backward goes against the grain of how the Federal Circuit describes its approach to claim construction: the nature of the accused devices is not supposed to influence claim construction. []

First, courts can from time to time (preferably based on rules articulated in advance) flip from the default of fixing ideational meaning during claim construction to the exception of fixing denotational meaning.[194] This is precisely what the Federal Circuit did in *Schering Corp. v. Amgen Inc.*[195] The court denied a claim any depth at all by using a denotational theory of meaning to construe the term "IFN-α" to mean the set of molecules to which the term referred on the date the patent was filed.[196] By alternating between ideational and denotational constructions of different terms within the same claim, a court can selectively craft claim depth. To return to the graphic convention introduced in Part II.B, a court can modulate claim depth to achieve policy goals in the following manner:



Second, invoking their discretion under *Phillips*, courts can rely more heavily on the use of language in the specification when construing the meaning of claim language.[197] When courts construe claim terms in light of the specification, they adjust patent breadth with the goal of preventing the reach of literal claim scope into after-arising technology. Graphically illustrated, they do this:



[194] *See* Collins, *Things and Meanings*, *supra* note [], at 553–58 (noting the possibility of using the distinction between ideational and denotational meaning as a policy lever).

[195] Schering Corp. v. Amgen Inc., 222 F.3d 1347 (Fed. Cir. 2000). *See also* Collins, *Things and Meanings*, *supra* note [], at 552–53.

[196] *Schering*, 222 F.3d at 1353–54.

[197] This scenario demonstrates that patent breadth is important to consider when assessing the reach of literal claim scope into after-arising technology. *See supra* Part I.B.

2. The Disclosure Doctrines

Unless every patent claim with depth is to be held invalid, the revised, ideational theory of the peripheral claim must definitively rule out the "full scope" interpretation of the enablement doctrine and its parallel, strict interpretation of the written description doctrine.[198] There are, however, two models for the disclosure doctrine that are waiting in the wings to replace the "full scope" interpretation.

Early, pre-Federal-Circuit enablement cases state that there need only be a "reasonable correlation" between the disclosure and the claims.[199] A similar reasonableness approach is imaginable for written description, too.[200] This fuzzy approach is attractive if the scope-regulating disclosure doctrines are to be attentive to the distinction between classic and orthogonal improvements. The presence of orthogonal improvements within literal claim scope should never raise enablement or written description problems; the patentee's failure to teach the PHOSITA how to make an orthogonal improvement or to demonstrate "invention" or "possession" of an orthogonal improvement to the PHOSTIA should be irrelevant to the commensurability analysis.[201] In contrast, the presence of classic improvements within literal claim scope should be considered in the commensurability analysis.[202] This distinction could be included as a factor in a reasonable-commensurability inquiry.

Alternatively, there is precedent in the enablement doctrine for exempting all after-arising technology from the commensurability requirement.[203] Perhaps, all technology known to the PHOSITA at the time of filing must be enabled, but any technology of which the PHOSITA becomes aware only after the filing date need not be enabled.[204] This categorical approach would make the disclosure doctrines consistent

[198] *See supra* text accompanying notes [Part III.A].

[199] *In re* Vaeck, 947 F.2d 488, 495 (Fed. Cir. 1991); *In re* Fisher, 427 F.2d 833, 839 (C.C.P.A. 1970).

[200] [cites]

[201] *See supra* [Part II.D].

[202] *See supra* [Part II.D].

[203] *In re* Hogan, 559 F.2d 595 (C.C.P.A. 1977). For an in-depth history of *Hogan*, see Adams, *supra* note [], at 82–102. To date, there is no parallel to this categorical exemption of after-arising technology from enablement in the written description doctrine. Chiron Corp. v. Genentech, Inc., 363 F.3d 1247 (Fed. Cir. 2004).

[204] Robert Merges has referred to this rule as the *temporal paradox* of enablement and written description: the claim must be fully enabled at the time of filing, but the meaning of claim language is not fixed, and is thus able to grow, until the date of infringement. Robert P. Merges, *Rent Control in the Patent District: Observations on the Grady-Alexander Thesis*, 78 VA. L. REV. 359, 379 n.73 (1992). Merges' interpretation of claim construction doctrine conflicts with the Federal Circuit's black-letter rule that the meaning of claim language is fixed on the date of filing. *See supra* text accompanying note []. For this reason, the

with the revised, ideational theory of peripheral claims, but it would not allow the courts to use the disclosure doctrines to sanction greater patent depth for orthogonal improvements and less patent depth for classic improvements.[205]

3. Means-Plus-Function Claiming

Means-plus-function claims are a type of patent claim for which courts use special, statutorily specified rules of claim construction.[206] If an inventor claims her invention as a "means for" performing a specified function, the claim does not literally encompass any and all means of performing that function. Rather, the claim encompasses only the means actually disclosed in the patent specification and its equivalents.[207] Means-plus-function claims provide a way for patent applicants to opt out of a peripheral claiming regime and into a variant of a central claiming regime by using language that invokes the means-plus-function rules. [208] Usually, improvement under a central claiming regime does not raise any special issue, as there is no attempt to fix any aspect of claim scope *ex ante*.[209] However, the Federal Circuit has expressly held that the literal scope of a means-plus-function claim cannot encompass after-arising technology such as improvements.[210]

Orthogonal improvements undermine the Federal Circuit's assertion that means-plus-function claims cannot encompass after-arising technology. Orthogonal improvements do—and should—routinely infringe means-plus-function claims, contrary to the purported prohibition on literal infringement of after-arising technology. Software

---

paradox of enablement and claim construction is, to the extent that it exists, a *meaning paradox*, not a temporal paradox: the full scope of claim language denotationally construed on the date of filing must be enabled, but the full scope of the claim language ideationally construed on the date of filing is dispositive in infringement. Kevin Emerson Collins, *Enabling After-Arising Technology*, 34 J. CORP. L. 1083, 1098–1105 (2009). In either situation, however, it is necessary to have a foreseeability requirement: technologies that are highly desired, yet that cannot be made, on the date of filing must be enabled. Plant Genetic Sys. v. DeKalb Genetics Corp., 315 F.3d 1335 (Fed. Cir. 2003). *See* Collins, *supra*, at 1098–1105. Otherwise, as Jeff Lefstin has aptly stated, using the temporal paradox to structure the enablement doctrine is tantamount to reducing enablement to "the tautological proposition that the inventor need not enable technology that is not enabled." Lefstin, *supra* note [], at 1174, n.100.

[205] However, if courts can perform this task using claim construction doctrine, *see supra* Part III.C.1, then perhaps they do not need disclosure doctrines that can also be used to perform the same task.

[206] 35 U.S.C. § 116, ¶ 6 (2006).

[207] *Id.*

[208] Burk & Lemley, *Sign Posts*, *supra* note [], at []; Fromer, *supra* note [], at [].

[209] Fromer, *supra* note [], at 770–71.

[210] Chiuminatta Concrete Concepts, Inc., v. Cardinal Indus., Inc., 145 F.3d 1303 (1998).

"apparatus" claims are commonly drafted in means-plus-function format,[211] yet there has never been a litigated case in which the after-arising nature of the hardware executing the newly invented software was raised as a defense to literal infringement. Similarly, mechanical devices are often claimed in means-plus-function language,[212] yet the after-arising nature of a material in an accused device has never been raised as a non-infringement defense. The absence of any litigation on these points strongly suggests that means-plus-function claims routinely encompass orthogonal improvements, despite their status as after-arising technologies.

Fed Cir doctrine on means-plus-function claiming shows just how thoroughly overlooked orthogonal improvements actually are on two levels—both how blind the blind spot is and how blind the court is of the existence of the blind spot. Orthogonal improvements sometimes are not even recognized as after-arising technology.

4. The Doctrine of Equivalents

The DOE extends a patentee's protection beyond literal claim scope. If an allegedly infringing technology does not fall within a peripheral claim, the patentee may still obtain rights with respect to the technology if the technology is "equivalent" to the claimed technology.[213] The DOE layers the type of protection available under a central claiming regime on top of a peripheral claim.[214]

At first glance, the effect of the revised, ideational theory of the peripheral claim on the DOE might seem to be inconsequential. The DOE already allows a patentee to reach into after-arising improvements,[215] and the fact that literal claim scope can encompass after-arising improvements need not in principle alter the DOE's capacity to do so as well. What is challenged by the revised theory of peripheral claims, however, is the current presumption that all after-arising technology in general receives special consideration under the DOE. After-arising technology in general is today described as the "quintessential example of an enforceable equivalent."[216] The justification for this special status is that reasonable claim drafters cannot be responsible for crafting claim

[211] [cites]

[212] [cites]

[213] Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17 (1997).

[214] This is an oversimplification, given the formalistic rules that have been developed over the last several decades that restrict the reach of the DOE. *See, e.g.*, *id.* at 28–30 (concluding that, under the DOE, each claim limitation must have an equivalent element in the accused device).

[215] *Warner-Jenkinson*, 520 U.S. at 37.

[216] Smithkline Beecham Corp. v. Excel Pharm., Inc., 356 F.3d 1357, 1364 (Fed.Cir.2004).

language to encompass technologies that they cannot foresee.[217] Once the revised theory of peripheral claims is adopted and it is acknowledged that literal claim scope has depth, the special status for all after-arising technology under the DOE is revealed to be overbroad in relation to its justification. After-arising technology that falls outside of the breadth of a claim should make a court look more carefully at DOE infringement.[218] The patent drafter may not have had the language at her disposal to describe the after-arising technology (as language and technology often evolve hand in hand), or perhaps it is not reasonable to put a burden on her to foresee a technology that she could describe with language that is within her grasp. However, after-arising improvements that fall within the breadth of a claim but that do not fall within the depth of a claim should almost never be actionable equivalents under the DOE.[219] Here, non-infringement is in no way related to the inability of a patent drafter to use broader language in the peripheral claim at the time of filing. In sum, the distinction between claim breadth and depth should become an integral feature of the doctrine controlling the application of the DOE to after-arising technologies, including improvements.

E. Peripheral Claims and the "Spirit" or "Point of Novelty" of an Invention

In a central claiming regime, concepts like the "spirit," "essence," "core," "gist," "central features" or "point of novelty" of an invention are important determinants of claim scope. Central claims clearly rely on these concepts, as it is the similarity of the allegedly infringing technology to the spirit/essence/gist of the disclosed exemplar that is the metric by which infringement is determined.[220] Peripheral claims, however, are often defined in contradistinction to claims that employ these concepts to determine their scopes.[221] Because a peripheral claim is presumed to identify the full scope of the claim *ex ante*, the spirit/essence/gist of the invention is presumed not to be needed to determine claim scope.

The revised, ideational theory of peripheral claims, however, demonstrates that peripheral claims only identify the full scope of the claim *ex ante* in a contingent manner. They identify the full scope of claim terms *ex ante* in a sense in which "scope" is synonymous with "meaning," but they do not identify the full scope of claims *ex ante* in the sense of making enumerable all of the thing-types that are encompassed within the

[217] Sage Prods., Inc. v. Devon Indus., Inc., 126 F.3d 1420 (1997).

[218] *See supra* text accompanying notes [Part I.B] (discussing improvements that lie beyond claim breadth).

[219] *See supra* text accompanying notes [Part I.B] (discussing improvements that lie within claim breadth but perhaps beyond claim depth).

[220] [string cite]

[221] Burk & Lemley, *Sign Posts*, *supra* note [], at 1747 ("Whereas peripheral claiming purports to mark the outermost boundary of the patentee's claims, central claiming describes the core or gist of the patentee's contribution to technology.").

claim.[222] In turn, the need to make a distinction between classic and orthogonal improvements means that the advance that justifies the issuance of the patent—i.e., the spirit/essence/gist of the invention—remains relevant to the determination of claim scope even in the contemporary peripheral claiming regime. In fact, the "advance" that underlies the patented invention—a concept that plays a starring role in any attempt to distinguish classic and orthogonal improvements—is arguably nothing more than the gist or spirit of the patented invention. To the extent that it recognizes the distinction between orthogonal and classic improvements, the contemporary peripheral claiming regime is already both a peripheral claiming regime and a regime in which the spirit/essence/gist of the invention is relevant to the determination of claim scope.[223]

CONCLUSION

Doctrine and scholarship on the temporal reach of patent scope into improvement has to date been farsighted. It has identified and addressed one type of improvement—classic improvement—but it has failed to notice another empirically common type of improvement that is right under our noses—orthogonal improvement.

This Article argues that the oversight must be corrected because classic and orthogonal improvements raise distinct normative concerns. The temporal reach of patent scope into classic improvements is legitimately contested, but patents should routinely encompass orthogonal improvements. The need to give patent scope sufficient temporal reach to encompass orthogonal can most clearly be seen by framing the successive advances wound up in improvement stories as economic complements and the orthogonal improvement itself as an indivisible bundle of the advances.[224]

Because orthogonal improvements routinely fall within literal claim scope, this Article also undermines the fixation theory of peripheral claims under which literal claim scope has not depth and is therefore limited to the set of things disclosed to the PHOSITA on the date of filing. Not only does the expansion of the literal scope of a peripheral claim over time need to be recognized, but its selective expansion must be acknowledged. Patent doctrine should overtly recognize that the doctrine that courts use to craft both patent depth and breadth should be sensitive to the distinction between classic and orthogonal improvements.

---

[222] *See* Collins, *Things and Meanings*, *supra* note [], at 502–03 (distinguishing "thing-scope" and "meaning-scope").

[223] [Oskar Liivak, *Finding Invention*] If a peripheral claiming regime is defined in contradistinction to a central claiming regime in which the spirit/essence/gist of the invention is relevant to the determination of claim scope, then the contemporary claiming regime is already a hybrid of true peripheral and central claiming regimes. This type of hybridity, however, is different from those commonly recognized. *See, e.g.*, Burk & Lemley, *Sign Posts*, *supra* note [], at []; Fromer, *supra* note [], at [].

[224] *See supra* part III.D.3.