

LEGAL SETS

Jeremy N. Sheff†

In this Article, I propose that the practices of legal reasoning and analysis are helpfully understood as being primarily concerned not with rules or propositions, but with sets. This Article develops a formal model of the role of sets in the practices of legal actors in a common-law system defined by a recursive relationship between cases and rules. In doing so, it demonstrates how conceiving of legal doctrines as a universe of discourse comprising (sometimes nested or overlapping) sets of cases can clarify the logical structure that governs marginal cases and help organize the available options for resolving such cases according to their form. While many legal professionals may intuitively navigate this set-theoretic structure, the formal model of that structure has important implications for legal theory. In particular, it (1) generates a useful account of the relationships among rules, standards, and principles; (2) provides a novel set of tools for understanding the nature of precedent; and (3) illuminates an extra-linguistic dimension to the problem of judicial discretion. On the last point, I argue that discretion is not merely a product of the imperfect relationship between abstractions and reality, or between natural language and the world, but that it is instead an emergent property of the structure of legal practice: a structure composed of sets “all the way down.”

† Professor of Law & Faculty Director, Intellectual Property Law Center, St. John's University School of Law. I am grateful for comments and suggestions from Brian Bix, David Carlson, Marc DeGirolami, Greg Keating, Anita Krishnakumar, Sara Lawsky, Josh Sarnoff, and Larry Solum. All errors are the author's alone.

TABLE OF CONTENTS

INTRODUCTION	2031
I. NAÏVE SET THEORY: A PRIMER	2038
A. <i>Membership, Construction, Equality, and Inclusion</i>	2039
B. <i>Set Algebra</i>	2043
1. Union and Intersection; Disjointness and the Null Set.....	2043
2. Difference and Complementation.....	2046
II. SETS IN LEGAL REASONING AND ANALYSIS	2048
A. <i>Building Blocks: The Relationship Between Rules and Cases</i>	2049
B. <i>More Complex Doctrines and Their Set-Theoretic Relations</i>	2054
1. Union and Subtraction: Overlapping Rules.....	2054
2. Inclusion: Special Circumstances, Included Offenses, Remedies Enhancements.....	2057
3. Intersection: Legal Elements and Multi-Pronged Rules.....	2061
C. <i>Common-Law Rule-Building: Extensional Definition</i>	2065
III. FORMAL STRATEGIES FOR AVOIDING CONTRADICTION	2076
A. <i>Competing Predicates and Logical Contradictions: Outlining the Problem</i>	2077
B. <i>The Trump Card</i>	2083
C. <i>The No True Scotsman</i>	2088
D. <i>The Tertium Quid</i>	2099
E. <i>Summary</i>	2111
IV. IMPLICATIONS AND LIMITATIONS	2112
A. <i>What the Set-Theoretic Model Is, and What It Is Not</i>	2112
B. <i>Precision and Generality: Rules and Standards, Policies and Principles</i>	2116
C. <i>Dimensions of Discretion: Precedent</i>	2121
D. <i>Meta-Rules: Sets All the Way Down</i>	2125
CONCLUSION.....	2132

INTRODUCTION

Legal reasoning and analysis are helpfully understood as being primarily concerned, not with *rules*, or even with *propositions*, but with *sets*. Set-based logic permeates mature legal systems, and constitutes the deep structure connecting legal authorities to the behavior of legal actors. This is particularly so in common-law systems, where judicial opinions serve a dual function as both backward-looking dispositions and forward-looking authorities, setting up a recursive dynamic that can best be modeled using the logic of sets. In this Article, I will explain that logic and describe some of its implications.

Consider that in many areas of law, our adversarial legal system channels disputes toward binary choices: Does this complaint state a cause of action or not? Is this statute constitutional or not? Is this defendant liable or not? Is this claimed element of damages recoverable or not?¹ Even where a rule appears *not* to be framed in terms of binary outcomes, it is usually trivially easy to re-frame it in such a way: instead of asking “how broad is the plaintiff’s right?” we might instead ask “does the plaintiff have a right to x enforceable against this defendant?”² Indeed, this is an essential move in legal reasoning: the process of resolving legal disputes often consists of channeling amorphous, complex issues into a series of discrete binary questions winnowed and tested via the adversarial process.³ Reaching one of those binary outcomes typically depends on a judgment as to whether the facts of a case satisfy some test defined by a legal rule: Does this paragraph of the plaintiff’s complaint recite “mere conclusory statements”?⁴ Is this statute’s effect on private speech limited to the punishment of “fighting

¹ See RONALD DWORKIN, A MATTER OF PRINCIPLE 119–20 (1985).

² Cf. *generally* WESLEY NEWCOMB HOHFELD, FUNDAMENTAL LEGAL CONCEPTIONS AS APPLIED IN JUDICIAL REASONING: AND OTHER LEGAL ESSAYS (Walter Wheeler Cook ed., 1919).

³ See MARC O. DEGIROLAMI, THE TRAGEDY OF RELIGIOUS FREEDOM 223 n.24 (2013) (citing TIMOTHY ANDREW ORVILLE ENDICOTT, VAGUENESS IN LAW 72 (2000) (“It is a consistent feature of legal systems that legal institutions treat legal standards as if their application were bivalent.”)).

⁴ See *Ashcroft v. Iqbal*, 556 U.S. 662, 678–79 (2009).

words”?⁵ Would this liquidated damages award constitute a “penalty”?⁶ And so on.

My first claim is that all such legal rules can be understood to *define categories*, and the application of those rules to the facts of a particular case thus consists of a determination whether a particular state of the world falls within or without a relevant category. This first claim marks a minor departure from the sentential deontic models common in contemporary law and logic theory,⁷ but it has clear antecedents in Anglo-American legal theory. H. L. A. Hart noted that “[a]ll rules involve recognizing or classifying particular cases as instances of general terms.”⁸ Likewise, Frederick Schauer draws heavily on the notion of categories (and the related notion of *generalizations*) in his philosophical investigation of the nature of rules.⁹ Indeed, all analysis—legal or otherwise—may at some level be reduced to this type of “lumping” and “splitting.”¹⁰ But in this Article, my aim is to *formalize* the insight into a model of legal analysis, because such formalization yields new insights emergent from the model itself. As one prominent history of formal logic points out:

Formalization is a difficult and tricky business, but it serves a valuable purpose. It reveals structure and function in naked clarity, as

⁵ Compare, e.g., *R.A.V. v. City of St. Paul*, 505 U.S. 377, 386–96 (1992), with *Snyder v. Phelps*, 131 S. Ct. 1207, 1215 n.3 (2011); *id.* at 1226–27 (Alito, J., dissenting).

⁶ U.C.C. § 2-718 (AM. LAW INST. & UNIF. LAW COMM’N 2017); see, e.g., *Equitable Lumber Corp. v. IPA Land Dev. Corp.*, 344 N.E.2d 391, 397 (N.Y. 1976).

⁷ See generally PABLO. E. NAVARRO & JORGE L. RODRÍGUEZ, *DEONTIC LOGIC AND LEGAL SYSTEMS* (2014).

⁸ H. L. A. HART, *THE CONCEPT OF LAW* 123 (2d ed. 1994); see also *id.* at 124 (“[T]he law must predominantly, but by no means exclusively, refer to *classes* of person, and to *classes* of acts, things, and circumstances; and its successful operation over vast areas of social life depends on a widely diffused capacity to recognize particular acts, things, and circumstances as instances of the general classifications which the law makes.”) (emphasis in original).

⁹ See generally FREDERICK SCHAUER, *PLAYING BY THE RULES: A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND IN LIFE* 18–21 (1991).

¹⁰ Bradley C. Karkkainen, “*New Governance*” in *Legal Thought and in the World: Some Splitting as Antidote to Overzealous Lumping*, 89 MINN. L. REV. 471, 479 (2004) (“It is sometimes said that the two most basic intellectual moves are ‘lumping’ and ‘splitting’—that is, finding relevant common characteristics that allow us intelligently and usefully to group apparently distinct phenomena into a single category (‘lumping’), and finding relevant distinguishing characteristics that allow us intelligently and usefully to separate otherwise similar phenomena into distinct classes (‘splitting’).”).

does a cutaway working model of a machine. When a system has been formalized, the logical relations between . . . propositions are exposed to view; one is able to see the structural patterns of various “strings” of “meaningless” signs, how they hang together, how they are combined, how they nest in one another, and so on.¹¹

This type of formalization must be distinguished from the typical meaning of the term “formalism” in legal theory. Historically, “formalism” in Anglo-American jurisprudence is (fairly or not) a pejorative term: it is a foil for the distinctively American school of Legal Realism.¹² Even in its most sophisticated iterations, this concept of formalism is at bottom a *substantive* and indeed a *normative* construct: it distinguishes legitimate bases for judicial decision-making from illegitimate ones, and makes claims regarding the extent to which legal authorities determine adjudicative outcomes or constrain judicial behavior.¹³ In this Article, I will avoid most of the normative questions that divide realists and anti-realists in an attempt to clarify the formal *structure* of legal analysis in a model that has explanatory power *regardless* of one’s substantive views or normative commitments regarding the nature of adjudication. As we will see, however, this purely formal model lends some support to realist claims regarding the indeterminacy of legal rules and to positivist claims regarding the primacy of practice in determining what a society’s law is.

My formal model will be built upon a particular vocabulary. In philosophy, logic, and mathematics, category definition is the province of *set theory*. Thus, if categories truly do play an important role in legal analysis, the tools of set theory may fruitfully be applied to the relationship between legal rules and particular cases, and ultimately to the relationships among legal rules themselves. For these purposes, a “rule” is any legal directive, formulated at any level of precision or generality, that purports to direct the behavior of actors within a legal

¹¹ ERNEST NAGEL & JAMES R. NEWMAN, GÖDEL’S PROOF 26 (Douglas R. Hofstadter ed., rev. ed. 2001).

¹² Compare BRIAN Z. TAMANAHA, BEYOND THE FORMALIST-REALIST DIVIDE: THE ROLE OF POLITICS IN JUDGING (2010), with Brian Leiter, *Legal Formalism and Legal Realism: What Is the Issue?*, 16 LEGAL THEORY 111 (2010).

¹³ See generally Leiter, *supra* note 12.

system.¹⁴ A “case” is any state of the world that has generated a legal dispute subject to evaluation within that legal system. In this way of thinking, what lawyers describe as “application of law to facts” is best understood as evaluation of the set-theoretic concept of *membership* or *belonging*: whether a particular state of the world can be situated within a category defined by a relevant legal rule.

Starting with this foundational notion of the relationship between rules and cases, I construct in this Article a set-theoretic and purely formal model of the structure of legal doctrine in a common-law system. In doing so I will rely primarily (though not exclusively) on examples from my own fields of substantive expertise: property and intellectual property (IP) law. For example: whether or not a transferee of a leasehold interest can be held directly liable to the original lessor for failure to pay rent often depends on whether the transfer falls into the category of “assignment” or the category of “sublease.” Which of those two categories any particular transfer falls into typically depends, in turn, on whether the instrument of transfer provides for the original lessee to retain any part of the leasehold interest.¹⁵ This set of rules can be formalized and related to one another in an overarching structure via the set-theoretic relations of membership and inclusion—as I will illustrate below.¹⁶

Moreover, as we will see, in common-law systems this structure is *recursive*: rules inform the disposition of cases, while the dispositions of cases collectively and inductively inform the development of new rules over time. Conceiving of legal doctrine in terms of sets, rather than propositions, allows a clearer understanding of the dynamic process by which the *practices* of legal actors generate the development and modification of legal directives. And importantly, it reveals how certain

¹⁴ The word “rule” here is thus not intended to reflect the particular meaning of the word “rule” in contrast with the word “standard” as the dichotomy is typically framed in legal theory; I discuss the relationship between my set-theoretic model and the rules/standards dichotomy (including the dimensions of precision and generality) in Section IV.B, *infra*.

¹⁵ See, e.g., *Thomas v. United States*, 505 F.2d 1282, 1286 (Ct. Cl. 1974) (quoting *Haynes v. Eagle-Picher Co.*, 295 F.2d 761, 763 (10th Cir. 1961)); *Duck Creek Tire Serv., Inc. v. Goodyear Corners, L.C.*, 796 N.W.2d 886, 893 (Iowa 2011); *Hartman Ranch Co. v. Associated Oil Co.*, 73 P.2d 1163, 1168 (1937).

¹⁶ See *infra* Section II.A.

dynamics of doctrinal change emerge predictably from the *structure* of a common-law system, rather than the *substance* of particular rules.

Formalization along these lines is the key to my second claim: that understanding the structure of legal analysis in terms of set theory reveals subtle *formal* distinctions in the behavioral strategies legal actors may deploy when faced with underdeterminate (or contradictory) legal rules. Thus, regardless of their *substantive* conclusions regarding the outcome of a particular case, and indeed independently of the *justification* for that outcome, legal actors will often have the freedom to implement their conclusion in a number of *formally* distinguishable ways. Distinctions among these strategies map directly to distinctions between the set-theoretic relations of membership and inclusion. We will see that the choice of a formal strategy has important implications for the claims made by any particular act of advocacy or adjudication against the body of doctrine within which it is situated.

Understanding these formal distinctions in the behavior of legal actors leads to my third claim: that important objects of study and debate in substantive legal theory—such as the distinction between rules and standards, the nature of precedent, and the problem of judicial discretion—are emergent properties of the set-theoretic structure within which legal practice operates: a system of cases categorized according to multiple, overlapping rules. In short, the logical structure of the relationship between rules and cases is what *generates* some of our most persistent jurisprudential concerns. These concerns are emergent features of the system's most elementary structures.

To be clear: I intend to demonstrate that a set-theoretic model of the structure of legal systems is extremely useful in understanding those systems, but I do not claim—nor do I in fact believe—that set theory can provide a *complete* model of legal systems, nor that it can answer all the important questions in legal theory or jurisprudence. I am not here making a claim about what law *is* in any metaphysical sense, about what makes it law as opposed to something else, about the *appropriate* scope of precedent, about the *appropriate* degree of constraint or discretion for judges, about the sources of normative content that do or should guide judicial decision-making, about law's relation to morality or to tradition, or any similar jurisprudential concern. Nor do I intend to defend any claim about the *proper* interpretation of legal texts or any other prescriptive theory of adjudication. On the contrary, my point is merely that the logical structure of legal doctrine necessarily implies

certain *behavioral options* in the context of specific legal disputes, and that these options are in fact *independent* of any such substantive or normative concerns—that they emerge from the structure of the legal system itself.

These claims, though modest, fill an unfortunate gap in the literature. There has never been an effort to formalize legal reasoning along the lines I attempt in this Article. This may be because the theoretical tools I will rely on were just gaining traction in philosophy at roughly the same time the American Legal Realists launched their program against formalist models of adjudication,¹⁷ and have since been largely abandoned to mathematicians as twentieth-century philosophy took its linguistic turn.¹⁸ Today, the formal logic of legal systems is a subject that is mainly of interest to philosophers,¹⁹ to those interested in the possibility of representing legal reasoning using the tools of artificial

¹⁷ For example, when Russell and Whitehead's *Principia Mathematica* was first published, Jerome Frank was already a law student on his way to private practice. Compare Frank, *Jerome New*, FED. JUD. CTR., <https://www.fjc.gov/history/judges/frank-jerome-new> [<https://perma.cc/TC5Q-PX6K>] (last visited Mar. 26, 2019), with Andrew David Irvine, *Principia Mathematica*, STAN. ENCYCLOPEDIA PHIL., <https://plato.stanford.edu/archives/spr2015/entries/principia-mathematica> [<https://perma.cc/ZXN4-NA96>] (last updated Mar. 10, 2015). By the time the disciplines of philosophy and mathematics had sufficiently digested the *Principia* to generate Gödel's breakthroughs, Karl Llewellyn and Max Radin were busy advancing the Realist program from perches on prominent law school faculties. See generally, e.g., Karl N. Llewellyn, *A Realistic Jurisprudence—The Next Step*, 30 COLUM. L. REV. 431 (1930); Max Radin, *Statutory Interpretation*, 43 HARV. L. REV. 863 (1930). Indeed, when Radin refers to the “determinables” and “determinates” of W. E. Johnson's *Logic*, in his disquisition on statutory interpretation, he is unwittingly relying on a philosopher whose efforts were soon to be overtaken by the system of Russell and Whitehead (and by the increasingly formal logics that followed), in keeping with the more general movement of the study of logic from philosophers to mathematicians. See *id.* at 868–70; JOHN PASSMORE, *A HUNDRED YEARS OF PHILOSOPHY* 343–45, 394 (2d ed. 1966).

¹⁸ The “linguistic turn” is a term of art in intellectual history and refers to an increased focus in post-war analytic (i.e., Anglo-American) philosophy on the philosophy of language, for which first-order symbolic logic of the type relied on in this Article was quickly deemed unsuitable. The term itself was popularized in *THE LINGUISTIC TURN: ESSAYS IN PHILOSOPHICAL METHOD* (Richard M. Rorty ed., 1992). See generally Peter M.S. Hacker, *The Linguistic Turn in Analytic Philosophy*, in *THE OXFORD HANDBOOK OF THE HISTORY OF ANALYTIC PHILOSOPHY* 926 (Michael Beaney ed., 2013).

¹⁹ See generally, e.g., Carlos E. Alchourrón, *On Law and Logic*, 9 *RATIO JURIS* 331 (1996); Richard Holton, *Modeling Legal Rules*, in *PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW* 165 (Andrei Marmor & Scott Soames eds., 2011).

intelligence,²⁰ and to scholars in the code-based civil law tradition²¹—with some notable and worthy exceptions.²² But lawyers and common-law legal scholars can benefit from greater attention to the common law as a distinctive logical system.

One important area of application is in legal education. The model developed in this paper will be useful to legal educators and law students, as a guide to the types of analytical moves that are part of “thinking like a lawyer.” For law students who tend to think graphically or spatially, in particular, the familiar graphical representations of sets in Venn diagrams is likely to be a particular aid to understanding and mastery of legal habits of mind such as analogy and distinction, the extraction of rules from cases, and the flexible scope of rule-application. Second, the tools of set theory are also of use to the practicing lawyer, judge, and legal commentator, because they offer a more precise vocabulary for identifying and critiquing poor legal reasoning that the typical rhetorical approach may gloss over. Finally, the set-theoretic model I will develop here has implications for legal theory, where heated substantive debates often leave important *formal* ambiguities

²⁰ See generally, e.g., Carlos E. Alchourrón, *Philosophical Foundations of Deontic Logic and the Logic of Defeasible Conditionals*, in DEONTIC LOGIC IN COMPUTER SCIENCE: NORMATIVE SYSTEM SPECIFICATION 43 (John-Jules Ch. Meyer & Roel J. Wieringa eds., 1993); John F. Horty & Trevor J.M. Bench-Capon, *A Factor-Based Definition of Precedential Constraint*, 20 ARTIFICIAL INTELLIGENCE & L. 181 (2012); Henry Prakken, *Logical Tools for Modelling Legal Argument: A Study of Defeasible Reasoning in Law*, 64 STUDIA LOGICA: AN INT’L J. FOR SYMBOLIC LOGIC 143 (2000).

²¹ For example, one recent compilation of essays (in English) on legal logic included contributions from twenty-four authors, only six of whom are from common-law countries. THE LOGIC OF LEGAL REQUIREMENTS: ESSAYS ON DEFEASIBILITY, at xi (Jordi Ferrer Beltrán & Giovanni Battista Ratti eds., 2012). See also, e.g., NAVARRO & RODRIGUEZ, *supra* note 7 (Argentinian authors); Jaap Hage, *Law, Logic and Defeasibility*, 11 ARTIFICIAL INTELLIGENCE & L. 221 (2003) (Dutch author).

²² See generally Horty & Bench-Capon, *supra* note 20; Henry Prakken & Giovanni Sartor, *Modelling Reasoning with Precedents in a Formal Dialogue Game*, in JUDICIAL APPLICATIONS OF ARTIFICIAL INTELLIGENCE 127 (Giovanni Sartor & Karl Branting eds., 1998). One leading American scholar of legal logic, Sara Lawsky, not coincidentally works primarily in the heavily codified field of tax law. See generally, e.g., Sarah B. Lawsky, *A Logic for Statutes*, 21 FLA. TAX REV. 60 (2017). Other innovative scholarship has focused on the potential of non-classical logics to illuminate particular aspects of legal reasoning such as fact-finding. See generally, e.g., Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic*, in LAW AND THE NEW LOGICS 32 (H. Patrick Glenn & Lionel D. Smith eds., 2017); Vern R. Walker, *A Default-Logic Paradigm for Legal Fact-Finding*, 47 JURIMETRICS J. 193 (2007).

unexamined. Not only can the set-theoretic model help cut through those ambiguities, but it also shows how some of these *substantive* debates are actually generated by inescapable *structural* features of law as a social practice.

The Article proceeds in four parts. Part I provides a brief overview of naïve set theory for those unfamiliar with it, introducing the terminology and concepts that will be deployed in the analysis that follows; those who feel comfortable with the concepts and notation systems of set theory and predicate logic may skim or skip this section. Parts II and III build the scaffolding of a set-theoretic model of law. Part II demonstrates how the tools of set theory can be deployed to analyze the relationship between rules and cases, and how legal doctrines can be modeled using set-theoretic concepts. Part III demonstrates more complex interactions of rules and cases, focusing on “hard cases” in which multiple applicable legal rules appear to contradict one another. This Part provides a more thorough description of the strategies legal actors can use to resolve such doctrinal conflicts at various levels of formal structure. Part IV discusses some implications and limitations of a set-theoretic understanding of legal doctrine, including its interaction with other aspects of legal theory.

I. NAÏVE SET THEORY: A PRIMER

Most people who have any degree of legal education will have at least a passing familiarity with the concept of sets—if for no other reason than because they are a staple of primary and secondary school mathematics, where they are the stuff of hazily remembered Venn diagrams. Sets are a pillar of modern mathematics: Axiomatic set theory is the basis for rigorous definitions of numbers, for setting up the rules of arithmetic and higher mathematics, for describing the nature of mathematical functions and geometry, and for constructing the types of abstract analyses of which modern mathematical proofs are made.²³ Fortunately, we require a far lesser degree of rigor to make sets useful for the analysis of legal doctrine. We will confine ourselves here to so-

²³ See generally JOSÉ FERREIRÓS, Labyrinth of Thought: A History of Set Theory and Its Role in Modern Mathematics (1999).

called “naïve” set theory of the type that one might encounter in an undergraduate, or even a high school, course—though some useful axioms will be introduced in a very informal way where appropriate. This Part also makes use of the standard notation system of predicate logic, familiarity with which is assumed.

A. *Membership, Construction, Equality, and Inclusion*

A “set” was defined by Georg Cantor as “any collection into a whole M of definite and separate objects m of our intuition or our thought. These objects are called the ‘elements’ of M .”²⁴ The relationship between a set and its elements—the relation of *belonging* or *membership*—is the most primitive relation in set theory,²⁵ and is the most important aspect of the theory for our purposes. For any given set A and one of its members x , we may say: “ x is an element of A ” or, in formal notation:

$$x \in A$$

Using this basic relation of membership, we can “construct” sets using at least two strategies. First, we can define a set by simply listing its elements. For example, as of this writing the set J of all active United States Supreme Court Justices is a finite set consisting of the following nine elements: Chief Justice Roberts, Justice Thomas, Justice Ginsburg, Justice Breyer, Justice Alito, Justice Sotomayor, Justice Kagan, Justice Gorsuch, and Justice Kavanaugh. We can formally define such a finite set *extensionally*: that is, we can define the set according to its elements merely by listing each and every one of those elements, i.e.:

²⁴ GEORG CANTOR, CONTRIBUTIONS TO THE FOUNDING OF THE THEORY OF TRANSFINITE NUMBERS 85 (Philip E. B. Jourdain trans., 1915). This definition was later found to be insufficient to ground a coherent set theory free of logical paradoxes (*see* note 28, *infra*), causing thinkers such as Zermelo, Russell, Quine, von Neumann, Gödel, Fraenkel, and others to refine the theory of sets to avoid such paradoxes. These refinements are the basis of axiomatic set theory, which is of great interest to mathematicians, computer scientists, and logicians, but of little use to the present project. A. A. FRAENKEL ET AL., FOUNDATIONS OF SET THEORY 22–28 (J. Barwise et al. eds, 2d rev. ed. 1973).

²⁵ PAUL R. HALMOS, NAIVE SET THEORY 2 (1960).

$$J = \{\text{Roberts, Thomas, Ginsburg, Breyer, Alito, Sotomayor, Kagan, Gorsuch, Kavanaugh}\}$$

This method of construction gives us occasion to invoke an important axiom: the *axiom of extensionality*, which provides the basis for the definition of the *equality* relation between sets. The axiom of extensionality specifies that sets are defined according to their *extension* (i.e., the full list of their elements), and therefore *two sets are equal to each other if and only if they have exactly the same elements*.²⁶

Extensional definition is obviously cumbersome for all but the smallest sets, and it is not especially informative regarding the qualities of the sets so defined. A more useful and parsimonious construction of a set of interest would invoke some common characteristic(s) that the elements of the set share, which distinguishes them from all other things we might have in mind that are not elements of the set. Here, an *intensional* definition of the set of interest—call this set J_I —is needed. To build such a definition, we require some *predicate statement* concerning members of our set that identifies them and distinguishes them from other objects. Our earlier casual definition, “the set of all active United States Supreme Court Justices,” suggests a useful predicate. We can define J_I as the set of all objects x such that x is an active United States Supreme Court Justice. In formal notation:

$$J_I = \{x: x \text{ is an active United States Supreme Court Justice}\}$$

Or, alternatively, using a predicate symbol to express our intensional definition:

$$P(x) = x \text{ is an active United States Supreme Court Justice}$$

$$J_I = \{x: P(x)\}$$

Note that once this predicate is *evaluated* against objects in the world and the resulting list of elements of J_I is fully extended, those elements will be exactly the same as the previously enumerated elements of J according to our extensional definition of J . Therefore, according to our definition of the equality relation between sets pursuant to the axiom of extensionality:

²⁶ FRAENKEL ET AL., *supra* note 24, at 22–28. Formally, using the notation system of first-order logic, we may state the axiom of extensionality thus:

$$\forall A \forall B [\forall x(x \in A \leftrightarrow x \in B) \rightarrow A=B]$$

$$J_1 = J$$

Some important features of our methods for constructing these sets bear mention. First, it should be apparent that our intensional definition of J_1 implies many other definitional statements that might be useful for distinguishing objects of study from one another. For example, the fact that x is a Supreme Court Justice implies that x is a human being, that x is a judge, that x is a federal judge, that x is an Article III federal judge. We may or may not need to make these implicit definitions explicit, depending on what the purpose of our analysis is. If we are interested in understanding the set of Supreme Court Justices in relation to other Article III judges, we might explicitly confine our analysis to Article III judges and define our sets as collections of objects from that larger group. For any given analysis, we can call the group of all objects that are candidates for inclusion in or exclusion from sets the *universe of discourse*.²⁷

Second, it should be clear that this understanding of the universe of discourse also defines a set. That is, the universe of discourse is a set of all objects that may come under consideration in a given analysis.²⁸ By

²⁷ GEORGE BOOLE, AN INVESTIGATION OF THE LAWS OF THOUGHT, ON WHICH ARE FOUNDED THE MATHEMATICAL THEORIES OF LOGIC AND PROBABILITIES 42 (1854) (“In every discourse, whether of the mind conversing with its own thoughts, or of the individual in his intercourse with others, there is an assumed or expressed limit within which the subjects of its operation are confined. . . . Now, whatever may be the extent of the field within which all the objects of our discourse are found, that field may properly be termed the universe of discourse.”).

²⁸ Understood at its most capacious and abstract, a universal set can pose fundamental challenges to the coherence of set theory. For example, if defined as “the set of all possible things,” U also by definition contains all sets containing members of U , and also contains U itself. This inherent self-reference leads to the potential for logical paradoxes, or “antinomies,” that arise from infinite recursion. The most famous of these is Russell’s Paradox, which posits the existence of a set composed of all sets that are not members of themselves. Such a set would be a member of itself if and only if it were not a member of itself—a logical contradiction. See BERTRAND RUSSELL, THE PRINCIPLES OF MATHEMATICS 101–03 (1996). But nothing in naïve set theory would seem to preclude an intensional definition of a set based on such an infinitely recursive predicate (i.e., that the set is a member of itself). The decline of Cantor’s naïve set theory, and the rise in the twentieth century of axiomatic theories that disclaim the existence of any truly Universal set, is a product of the effort to avoid this type of paradox.

Fortunately, for our purposes, we are not attempting to build an internally consistent system of logic or provide an axiomatic framework for all of mathematics, and we are not interested in plumbing the nature of infinity or of the set of all possible sets. We are instead merely using the language of sets as a convenient shorthand for the logical relationships

convention, we will call this universal set U . So, for our sets J and J_1 above, the universal set U might be defined as the set of all human beings, or of all judges, or of all federal judges, or of all Article III federal judges, depending on what universe of discourse we are trying to make meaningful statements about.

This leads us to a third important relation in naïve set theory: the relation of *inclusion*, which is based on the concept of the *subset*, indicated by the symbol \subseteq . Let us say—by way of definition—that a set A is a *subset* of another set B where each and every element of A is also an element of B . We may also say, conversely, that the set B *includes* its subset A .²⁹

Under this definition of inclusion, B may well have elements *in addition to* the elements of A , but it need not. Sets that are equal to one another are also subsets of one another, and every set is a subset of itself.³⁰ Furthermore, it should be apparent that the relation of inclusion is *transitive*: that is, if A is a subset of B and B is a subset of C , then A is necessarily a subset of C .³¹ Finally, it should be apparent that, whatever our universe of discourse, any set we define within that universe will by definition be *included in*, and therefore a *subset of*, the universal set U .

We can now move from the fundamental relations of membership, equality, and inclusion to some other set theoretical concepts of interest to the current project. Chief among these are *union*, *intersection*,

between certain clearly-defined objects of our interest and study, and the nature of our inquiry is such that we can define those objects and relationships in such a way as to avoid these paradoxes. In technical terms, we will take care to avoid any *impredicative* definitions of our sets, including our universal sets. See *id.*; see also FRAENKEL ET AL., *supra* note 24, at 38 (“A definition of a set is called impredicative if it contains a reference to a totality to which the set itself belongs.”). We will leave the formal objections to naïve set theory to the philosophers and mathematicians.

²⁹ FRAENKEL ET AL., *supra* note 24, at 26. Formally, the relation of inclusion is defined as follows:

$$\forall A \forall B [\forall x(x \in A \rightarrow x \in B) \rightarrow A \subseteq B]$$

³⁰ *Id.* A subset that is not equal to the set that includes it is sometimes referred to as a *proper subset*, but the distinction between a subset and a proper subset is not particularly useful to this Article.

³¹ That is, each and every element of A is an element of B , and each and every element of B is an element of C , meaning that each and every element of A must also be an element of C . Formally:

$$[(A \subseteq B) \wedge (B \subseteq C)] \rightarrow (A \subseteq C)$$

difference, and complementation; analysis of these operations will yield an understanding of *disjointness*, and of the *null set*.

B. Set Algebra

Sets can be the subject of mathematical and logical operations that make them useful analytical tools. Among these, the more interesting for our purposes have to do with the relationships between sets.

1. Union and Intersection; Disjointness and the Null Set

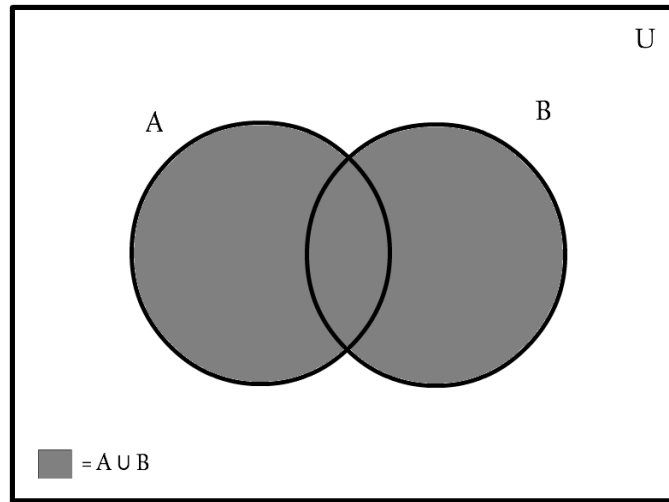
The *union* of two sets A and B , indicated by the symbol \cup , is defined as the set (call it C) containing all objects that are elements of *either A or B or both*.³² The union operation may be analogized to addition in arithmetic, or to disjunction in sentential logic. We may illustrate the concept of a union of sets graphically, using a Venn diagram representing the set C as a union of sets A and B .³³

Figure 1

³² ABRAHAM A. FRAENKEL, ABSTRACT SET THEORY 18–20 (3d rev. ed. 1966).

³³ We may also define the union function formally:

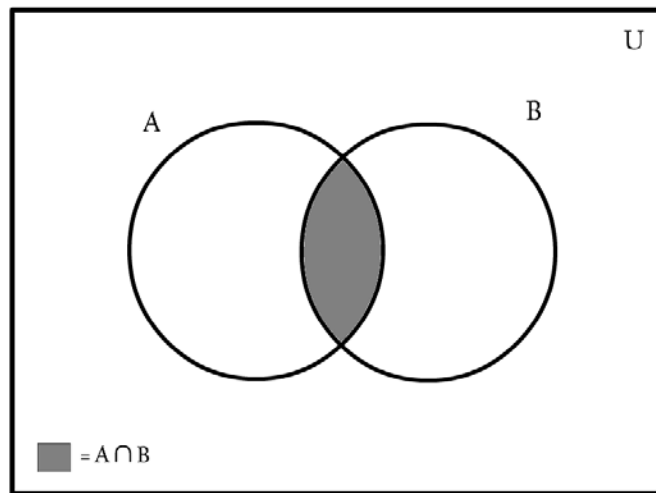
$$C = A \cup B \rightarrow C = \{x : (x \in A) \vee (x \in B)\}$$



Note that any set constructed as the union of two or more other sets necessarily *includes* those other sets; i.e., the sets that are added together are each subsets of the sum of those sets.

The *intersection* of two sets A and B , indicated by the symbol \cap , is defined as the set (call it I) containing all objects that are elements of *both* A and B .³⁴ It is thus analogous to conjunction in sentential logic. Again, the intersection of two sets can be represented graphically:³⁵

Figure 2



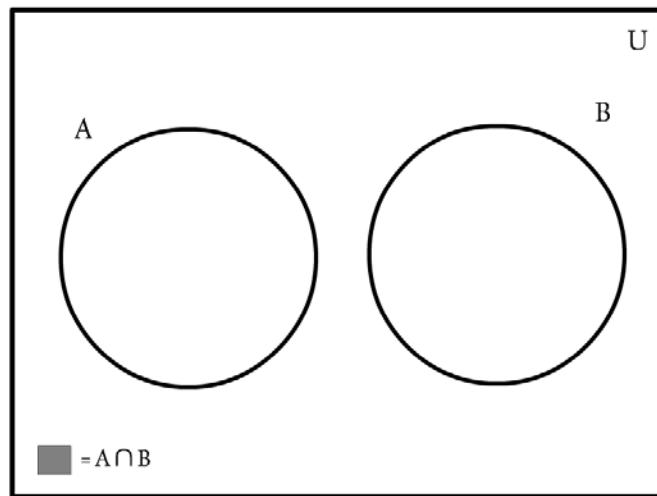
³⁴ FRAENKEL, *supra* note 32.

³⁵ Again, the intersection function can also be defined formally:

$$I = A \cap B \rightarrow I = \{x : (x \in A) \wedge (x \in B)\}$$

Note that nothing about the nature of sets as we have defined them up to this point requires our intersection set I to have any objects in it. Should it be the case that A and B have no elements in common, the set I will have no elements:

Figure 3



This set with no elements is unique: By the axiom of extensionality all sets with the same elements are equivalent, and any set with no elements is therefore equivalent to every other set with no elements. This unique set with no elements is called the *null set* or *empty set*, and is denoted by the symbol \emptyset .³⁶ Where, as in Figure 3, the intersection of two sets A and B is the null set (i.e., where A and B have no members in common), A and B are said to be *disjoint sets*.³⁷

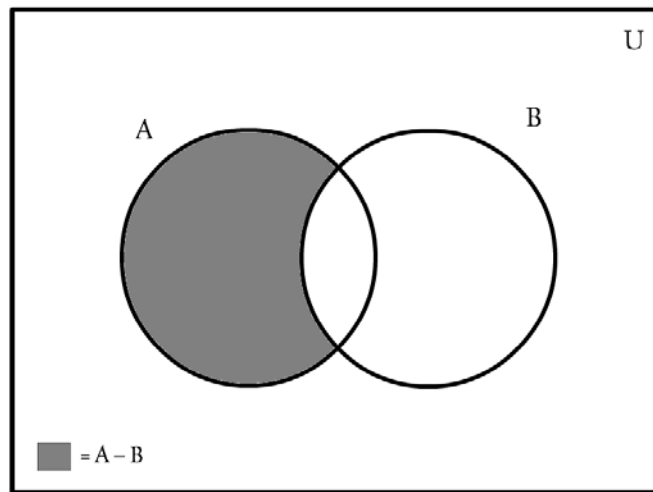
³⁶ FRAENKEL, *supra* note 32 at 16–18. Given our definition of subsets, it follows that the null set is a subset of every set, including itself. *Id.*

³⁷ FRAENKEL ET AL., *supra* note 24, at 30.

2. Difference and Complementation

The *difference* between two sets, $A - B$, is defined as the set of those members of A that are not members of B .³⁸ As a matter of sentential logic, it is analogous to conjunction with a negated premise. We can represent this definition of set difference graphically:³⁹

Figure 4



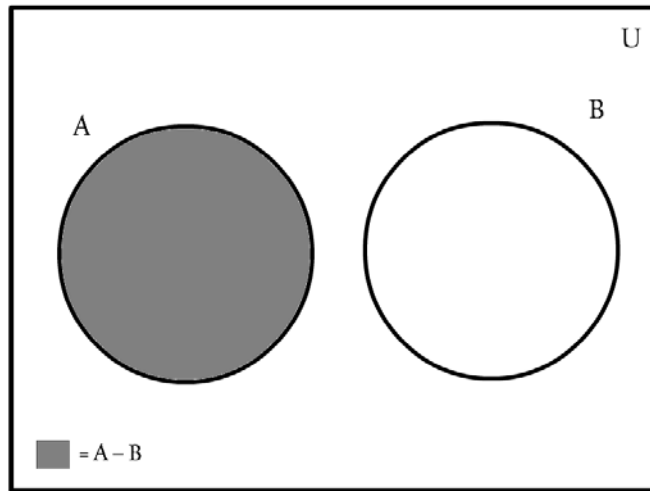
Of course, it is possible that A and B will be disjoint sets, in which case the difference $A - B$ will simply be equivalent to A :

³⁸ FRAENKEL, *supra* note 32, at 22–23.

³⁹ Formally, difference is defined as follows:

$$[C = A - B] \rightarrow [C = \{x : (x \in A) \wedge (x \notin B)\}]$$

Figure 5



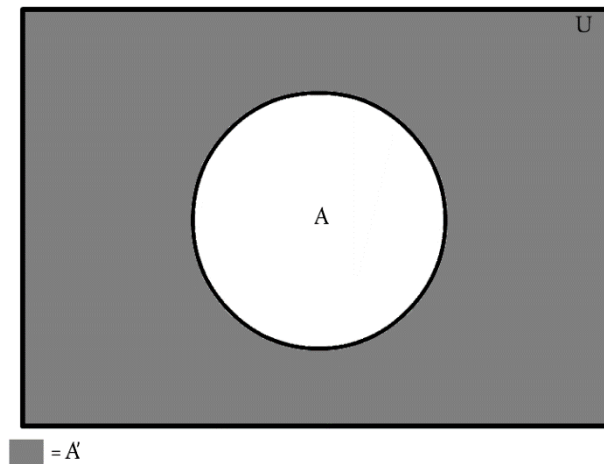
Finally, where the universe of discourse is suitably specified, any set A defined within the universal set U may be said to have a *complement*, which is defined as the difference between the universal set and the specified set, or $U - A$.⁴⁰ Put differently, this complement is the set of all objects in the universe of discourse that are *not* elements of A . The complement of the set A is designated as A' , such that:

$$A' = U - A$$

Again, we can represent this definition of complementarity graphically:

⁴⁰ HALMOS, *supra* note 25, at 17–18. It is precisely because of the paradoxes discussed in note 28, *supra*, that complementation (like the universal set) is generally not well defined in axiomatic set theory. FRAENKEL ET AL., *supra* note 24, at 40–41.

Figure 6



Having defined some basic tools of naïve set theory, we are now in a position to apply them to legal problems.

II. SETS IN LEGAL REASONING AND ANALYSIS

We can understand legal rules to provide *intensional* definitions of sets. Such intensional definitions may operate either by *inclusion* or by *exclusion*—some rules tell us what is *in* a category, others tell us what is *not* in a category. For example, the category of legal infants is often framed in an *inclusionary* way: it consists of all natural persons who are younger than a certain age.⁴¹ But the category of persons with legal capacity to sue or be sued is usually framed in an *exclusionary* way: it

⁴¹ See, e.g., CONN. GEN. STAT. ANN. § 1-1d (West 2019) (“[T]he terms ‘minor’, ‘infant’ and ‘infancy’ shall be deemed to refer to a person under the age of eighteen years”); but see KAN. STAT. ANN. § 38-101 (West 2019) (“The period of minority extends in all persons to the age of eighteen (18) years, except that every person sixteen (16) years of age or over who is or has been married shall be considered of the age of majority in all matters relating to contracts, property rights, liabilities and the capacity to sue and be sued.”). The fact that this category could also be framed in exclusionary terms is an example of the binary nature of so many legal doctrines—an issue which will be developed more fully below. See, e.g., 23 PA. STAT. AND CONS. STAT. ANN. § 5101 (West 2019) (“Except where otherwise provided or prescribed by law, an individual 18 years of age and older shall be deemed an adult”).

consists of all persons who are *not* infants, insane, incapacitated, or subject to some other legal disability.⁴²

On this understanding of the nature of legal rules, the application of law to facts—that inescapable and inescapably contested analytical process at the heart of legal systems—can be formally modeled as the evaluation of the fundamental set-theoretic relation of *membership*. Cases—not in the sense of “judicial opinions” or even “matters in litigation” but in the sense of “states of the world that have generated disputes subject to resolution by the legal system”—can be understood as objects that may or may not be elements of the sets intensionally defined by our legal rules. Other aspects of legal reasoning may similarly be modeled in set-theoretic terms. Rules may be combined (or reconciled) with one another using set-theoretic concepts such as union, subtraction, intersection, and inclusion. Finally, in the quintessential dynamic of common-law systems, judicial resolution of cases can, over time, provide *extensional* definitions of new legal sets, and new legal rules (i.e., *intensional* definitions) can be inferred by grouping judicial decisions according to their outcomes and attempting to formulate predicates that describe the states of the world—the cases—that generated those similar outcomes. In this Part, I will illustrate the use of set-theoretic concepts to model these basic features of legal reasoning and analysis.

A. *Building Blocks: The Relationship Between Rules and Cases*

Recall that the most fundamental relation in set theory is the relationship between sets and elements (the relation of *membership*). I propose that the analogue to this relationship in common-law legal systems is the relationship between rules and cases. As every first-year law student learns, the real work of legal analysis comes in applying a rule to a set of facts that has generated a legal dispute—to a *case*. Consider the example we introduced in the Introduction to this Article: the distinction between an assignment and a sublease.⁴³ Where a

⁴² 67A C.J.S. *Parties* § 10 (West 2019) (“A want of capacity to sue exists where there is some legal disability, such as infancy, lunacy, idiocy, coverture, want of authority, or a want of title in plaintiff in the character in which he or she sues.”).

⁴³ See *supra* note 15 and accompanying text.

transferee of some interest under a lease breaches some covenant of the lease agreement, the landlord under the primary lease may attempt to obtain monetary relief (as distinct from possession) directly against the breaching transferee. Whether such direct relief is available depends on whether the primary landlord and the transferee are “in privity,” and this, in turn, depends on whether the transfer at issue is categorized as a “sublease” or an “assignment.” Simplified somewhat, an assignment puts the original lessor and the transferee in privity of estate (and thus makes the transferee directly answerable to the original lessor in an action for damages); a sublease does not.⁴⁴ The hornbook rule for distinguishing among these two types of transfer is as follows:

An assignment of a term for years occurs where the lessee transfers his or her entire interest therein, for the unexpired remainder of the term created by the lease, without retaining any reversionary interest. . . . A sublease occurs where a lessee underlets the premises or a part thereof to a third person for a period less than the lessee’s term. If the lessee reserves a reversionary interest in the term, it constitutes a sublease, no matter how small the reversion and regardless of the form of the instrument.⁴⁵

Thus, in any given action for damages brought by a landlord against a transferee of a leasehold interest, the possible outcomes are—at least at one level—binary: the landlord-plaintiff either can recover or cannot recover against the transferee-defendant. Which of these two binary outcomes will result from any particular set of facts depends on categorization of the instrument by which the particular transferee acquired her interest as either an assignment or a sublease. In short, the state of the world that generated the dispute must be placed into one of the categories defined by the hornbook rule.

⁴⁴ See 52 C.J.S. *Landlord & Tenant* § 62 (West 2019) (“Since there is neither privity of contract nor privity of estate between the two, a subtenant normally owes no responsibility to the original lessor, and the original lessor has no direct action with respect to the covenants in the original lease as against the sublessee.”); 52 C.J.S. *Landlord & Tenant* § 54 (West 2019) (“When a lease is transferred by assignment, privity of estate ends between the lessor and lessee and is created between the lessor and the assignee, who becomes bound by covenants running with the land.”).

⁴⁵ 52 C.J.S. *Landlord & Tenant* § 43 (West 2019).

Some examples clarify this relationship between rules and cases. In *Berg v. Ridgway*,⁴⁶ the plaintiff landlords sued the third successive transferee of a ground lease for unpaid sums due under the primary ground lease agreement with the original tenant. The defendant had taken possession of the leased premises pursuant to a partnership dissolution agreement whereby the partnership—itsself a prior assignee of the lease—agreed to “transfer, sell, assign and convey all right title and interest . . . to the” defendant, who agreed to “purchase[] and accept[] such transfer, assignment and conveyance.”⁴⁷ The defendant claimed that he could not be held directly liable, based on language in the original lease governing the liability of subtenants. But because the dissolved partnership had retained no interest in the leased premises, the court held that the transfer to the defendant was an assignment, not a sublease, making the defendant directly liable to the original lessor for performance of the lease covenants (and thus for damages) notwithstanding the language in the primary lease.⁴⁸

By way of contrast, in *Dunlap v. Bullard*,⁴⁹ the defendant’s predecessor acquired an interest in an existing lease via an instrument that provided he was “to hold for a term equal to the whole of the unexpired term of the original lease,” but also providing that the transferor “might enter and take possession for breach of covenant,” and that the transferee “would quit and deliver up the premises to the lessor at the end of the term.”⁵⁰ The original lessor sued to recover for unpaid taxes that were the responsibility of the tenant under the primary lease. The Massachusetts Supreme Judicial Court ordered judgment for the defendant, reasoning that the reservation by the transferee of a right to retake possession was inconsistent with an assignment, insofar as it indicated “that the parties to this lease intended to create the relation of landlord and tenant between themselves”—that is, to create a sublease.⁵¹

⁴⁶ 140 N.W.2d 95 (Iowa 1966).

⁴⁷ *Id.* at 97.

⁴⁸ *Id.* at 100.

⁴⁹ 131 Mass. 161 (1881).

⁵⁰ *Id.* at 161.

⁵¹ *Id.* at 162–63.

As before, we can build a set-theoretic model of the doctrines in this area, either formulaically or graphically. First, we can conceive of the universe of discourse as divided into two complementary sets corresponding to the possible outcomes—intensionally defined as judgment for plaintiff and judgment for the defendant. Let π be the set of cases wherein a court enters judgment for the plaintiff and Δ be the set of cases wherein a court enters judgment for the defendant. Next, we can intensionally define two sets predicated by our hornbook rules: let A (for “assignment”) be the set of cases wherein the transferor has conveyed her entire interest in a lease, and let S (for “sublease”) be the set of cases wherein the transferor has conveyed less than her entire interest in a lease. Because our hornbook rule tells us that the predicates of the sets A and S determine the appropriate judicial resolution of a legal dispute, we know they must have some relation to our universe-exhausting complementary sets π and Δ . In particular, we know that any case within set S will—under the rules we have so far discussed—result in a judgment for the defendant, while any set of facts that falls within set A will result in a judgment for the plaintiff. Thus, any element of S will also be an element of Δ , while any element of A will also be an element of π . In other words, S must be a *subset* of Δ , and A must be a *subset* of π .

We cannot treat our legal-rule-defined sets as *equal* to our universe-exhausting complements for a very important reason: legal rules and inferences are generally defeasible.⁵² That is, it may be that there are *other* legal rules that might be brought to bear on a particular case that would play a more important role in determining which of our two binary outcomes is required in any particular case. We will consider that possibility—and the ways in which it may alter a set-theoretic model of a legal universe of discourse—in the following Section.

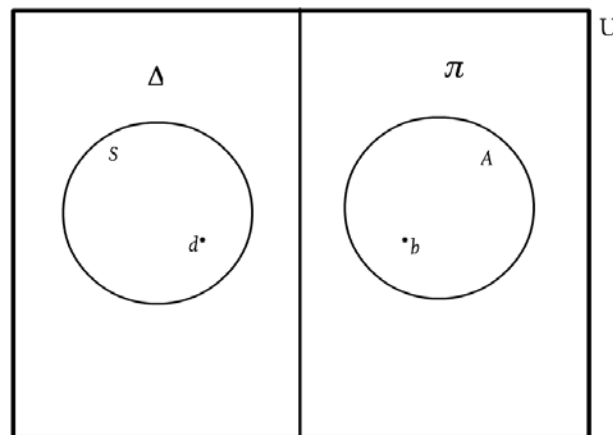
Before reaching that point, we need to incorporate actual cases into our model. If, as posited above, we should treat the states of the world

⁵² For theoretical approaches to the logical concept of defeasibility and its role in legal reasoning, see generally THE LOGIC OF LEGAL REQUIREMENTS: ESSAYS ON DEFEASIBILITY (Jordi Ferrer Beltrán & Giovanni Battista Ratti eds., 2012); Giovanni Sartor, *Defeasibility in Legal Reasoning*, in INFORMATICS AND THE FOUNDATIONS OF LEGAL REASONING 119 (1995); Henry Prakken & Giovanni Sartor, *The Three Faces of Defeasibility in the Law*, 17 *RATIO JURIS* 118 (2004); Richard H. S. Tur, *Defeasibilism*, 21 *OXFORD J. LEGAL STUD.* 355 (2001); John L. Pollock, *Defeasible Reasoning*, 11 *COGNITIVE SCI.* 481 (1987).

that generated the two judicial opinions discussed in this Section as objects that can be *elements* of sets in our universe of discourse, we must incorporate the cases into the model using the relation of *membership*. Let b stand for *Berg v. Ridgway*, and let d stand for *Dunlap v. Bullard*. We know that *Berg* was found to involve an assignment, and thus resulted in judgment for the plaintiff; we can therefore say that b is an element of set A and therefore—by inclusion—of set π . Similarly, we know that *Dunlap* was found to involve a sublease, and thus resulted in judgment for the defendant. We can therefore say that d is an element of set S and therefore—by inclusion—of set Δ .

This gives us sufficient material to construct the simplest possible model of a legal universe of discourse: a pair of rule-predicated sets mapping to a binary outcome that exhausts the possible resolution of cases that fall within the universe, with two exemplar cases as elements of our two rule-based sets. With all the pieces in place, we can represent this universe of discourse graphically:

Figure 7



If all legal doctrines were this simple, modeling them with set theory would seem to be more trouble than it is worth. But as any lawyer—or law student—knows (and as the rest of this Article will explore), legal rules typically interact in more complex ways. And set theory is a powerful tool for understanding that complexity by constructing a model of the relevant universe of discourse that organizes

the complex facts of real-world cases into comprehensible—if not entirely stable—categories.

B. *More Complex Doctrines and Their Set-Theoretic Relations*

1. Union and Subtraction: Overlapping Rules

We can begin to complicate the picture without leaving the law of subleases and assignments. Recall that in the previous Section we refrained from treating the legal-rule-defined sets S and A as *equal* to the universe-exhausting complements Δ and π because of the defeasibility of legal conclusions: It is possible that *other* legal rules might compel a particular outcome in a particular case *notwithstanding* the application of the rule regarding assignments and subleases. Our first example of the interaction of multiple legal rules involves one such additional rule: the rule regarding assumption of a primary lease by a transferee.

Recall that assignments render transferees liable to the original lessor while subleases do not because of the common-law requirement of privity—satisfied in the case of assignments by the notion of *privity of estate*.⁵³ But the common law developed another theory by which the lessor could hold the transferee directly liable for damages: if a transferee *agreed to be bound* to the covenants of a primary lease, that would make the lessor under that primary lease a third-party beneficiary of the transfer agreement, and thus create *privity of contract* between the lessor and the transferee.⁵⁴ Indeed, this theory was available—though superfluous—in *Berg v. Ridgway*.⁵⁵

⁵³ *Berg*, 140 N.W.2d at 99–100.

⁵⁴ The hornbooks recite this principle in the context of assignments, where it is far more commonly applied: “[I]n cases in which the assignee expressly assumes the lessee’s obligations under the lease, he or she may be bound to the lessor under privity of contract principles even where the lessor is not a party to the assignment.” 52 C.J.S. *Landlord & Tenant* § 54 (West 2019).

⁵⁵ 140 N.W.2d at 100 (“The trial court could find defendant had agreed to perform all the provisions of the lease during the period of his occupancy.”).

A third case provides an example. In *Hartman Ranch Co. v. Associated Oil Co.*,⁵⁶ Hartman Ranch executed an oil and gas lease to Joseph Dabney.⁵⁷ Dabney later executed an instrument—to which Hartman was not a party—transferring Dabney’s interest under the lease for the same term to Associated Oil.⁵⁸ Hartman subsequently alleged a breach by Associated of an implied covenant in the primary lease to Dabney and sued for damages in the form of unpaid royalties.⁵⁹ The transfer from Dabney to Associated “gave Dabney and his associates a right of re-entry for breach of any stipulation therein.”⁶⁰ According to the “Massachusetts Rule” applied in *Dunlap*, then, the transfer from Dabney to Associated Oil was a sublease rather than an assignment—Dabney had reserved a right to retake possession in the instrument of transfer.⁶¹ This might have resolved the case in favor of the transferee, Associated. But the transfer instrument also “contained an express promise whereby defendant assumed the parent Hartman lease,”⁶² and the California Supreme Court held that this covenant put Hartman and Associated into *privity of contract* through a third-party beneficiary relationship.⁶³ Thus, Hartman was permitted to recover damages against Associated on a contract—as opposed to a property—theory.⁶⁴

Hartman requires some revision of our set-theoretic model of the law of sublease and assignment. There are now *two* types of facts that will allow for a plaintiff’s recovery: *either* an assignment, *or* an assumption of the primary lease by the transferee. And whether or not the transferee has assumed the primary lease is a fact that is independent of the categorization of the transfer as a sublease or assignment: either form of transfer might or might not include such an assumption of covenants. We can model this slightly more complex universe of discourse by invoking additional set-theoretic concepts.

⁵⁶ 73 P.2d 1163 (Cal. 1937).

⁵⁷ *Id.* at 1166.

⁵⁸ *Id.*

⁵⁹ *Id.* at 1165–66.

⁶⁰ *Id.* at 1168.

⁶¹ *Id.*

⁶² *Id.*

⁶³ *Id.*

⁶⁴ *Id.* at 1169–71.

Let us identify a new set T , as the set of all cases in which a transferee of a leasehold interest assumes the covenants of the primary lease. Because assumption of the primary lease by the transferee provides an independent basis for a plaintiff's recovery, we know that T will be a subset of π . And because we know that *either* an assignment *or* an assumption of the primary lease will provide a basis for the plaintiff to recover, we can identify the relationship between sets A , T , and π with the logical operation of *disjunction*—which as previously discussed,⁶⁵ relates to the set-algebraic operation of union:

$$(A \cup T) \subseteq \pi$$

Indeed, assuming we have now identified and modeled all the legal rules applicable to the universe of discourse, and that these rules exhaust the relevant universe of discourse (which we will do here for simplicity's sake), we can simply posit:

$$\pi = (A \cup T)$$

Finally, we can identify all remaining cases in our simplified universe of discourse—which, given our assumptions, is simply the universe of all cases in which a landlord is seeking to recover damages directly from a transferee of a primary lease—as subleases in which the subtenant has not assumed the covenants of the primary lease. The logical form of this construction—conjunction with a negated premise—is, as we have previously noted,⁶⁶ analogous to the set-algebraic operation of *subtraction*:

$$\Delta = (S - T)$$

We can now re-evaluate our cases for membership in the sets intensionally defined by our three legal rules. We need to know whether each of our three cases involves a sublease or an assignment, and we also need to know whether the transferee assumed the primary lease in each case. Given the discussion above we know that *Dunlap* and *Hartman* involved subleases while *Berg* involved an assignment; we also know that the transferee assumed the primary lease in *Berg* and *Hartman* but not in *Dunlap*. Let h represent *Hartman*, just as b and d represent *Berg*

⁶⁵ See *supra* note 32 and accompanying text.

⁶⁶ See *supra* note 38 and accompanying text.

and *Dunlap*, respectively. We can say that b is an element of sets A and T but *not* of set S , that d is an element of set S but *not* of sets A or T , and that h is an element of sets S and T but *not* of set S . With these membership relations in place, and given the previously described relationships between the sets in our universe of discourse, we can now represent our universe of discourse graphically:

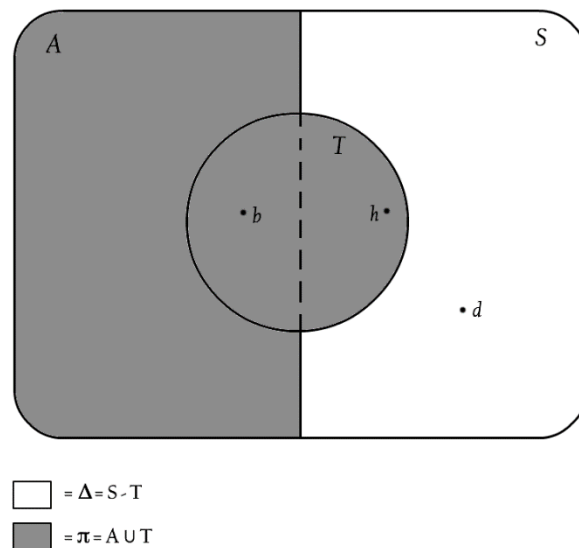


Figure 8

The use of union and subtraction to organize multiple applicable legal rules is the simplest means of resolving potential conflicts between such rules. But as Part III will demonstrate, it is not the only way of doing so. Moreover, the different strategies for resolving such conflicts are intimately related to the process of common-law doctrinal development. But before exploring these strategies, we must first examine the role of sets in other aspects of legal analysis.

2. Inclusion: Special Circumstances, Included Offenses, Remedies Enhancements

In building our first models of legal doctrine using set-theoretic tools we conceived of the sets predicated by legal rules as subsets of our

ultimate sets of interest: the complementary sets defining a binary outcome of adjudication. This is a key role for the relation of inclusion in set-theoretic models of legal doctrines, and one which we will examine more fully in the following Part. But subsets—and the relation of inclusion—can play other roles in basic legal analysis as well. The nesting of legal categories within one another, according to the relation of inclusion, is characteristic of all legal rules regarding special circumstances that select a specific outcome within a more general class of related outcomes. For example: in civil cases, remedies enhancements may be triggered in a subset of cases that entitle the plaintiff to some recovery.⁶⁷

Other legal doctrines that define questions of degree similarly rest on the logic of inclusion. Consider the grading of crimes under the Model Penal Code. We can take property crimes as an example. In general, “[a] person is guilty of theft if he unlawfully takes, or exercises unlawful control over, movable property of another with purpose to deprive him thereof.”⁶⁸ But once the elements of the crime of theft are established (an issue we will deal with in the next Section), the grade of the crime will depend on other additional facts that may or may not be proven in a particular case. Thus, theft of property with value of less than \$50, achieved without breach of fiduciary duty, threat, or taking from the person of the victim, is a petty misdemeanor. Theft of property with value exceeding \$500, or of a firearm, or of a motor-propelled vehicle is a third-degree felony. All other theft is a misdemeanor.⁶⁹ We can understand each of the grades of theft under the Model Penal Code as a *subset* of the crime of theft.

Again, we can express these relations both formally and graphically. Consider a universe of discourse consisting of states of the world in which a criminal defendant has committed the crime of theft. Within this universe of discourse, we can define *subsets* of our universal set according to the various facts that go not to the *commission* of the crime, but to its *degree*:

⁶⁷ See, e.g., 17 U.S.C. § 504(c) (2010) (providing statutory damages of up to \$30,000 per work for copyright infringement, but increasing this maximum to \$150,000 where the infringement is willful).

⁶⁸ MODEL PENAL CODE § 223.2(1) (AM. LAW. INST. 1962).

⁶⁹ MODEL PENAL CODE § 223.1(2)(a)–(b) (AM. LAW. INST. 1962).

$$A = \{x: x \text{ is a theft of property with value between \$50 and \$500}\}$$

$$B = \{x: x \text{ is a theft achieved by breach of fiduciary duty}\}$$

$$C = \{x: x \text{ is a theft achieved by threat}\}$$

$$D = \{x: x \text{ is a theft achieved by taking from the person of the victim}\}$$

$$E = \{x: x \text{ is a theft of property with value exceeding \$500}\}$$

$$F = \{x: x \text{ is a theft of a firearm}\}$$

$$G = \{x: x \text{ is a theft of a motor-propelled vehicle}\}$$

We can also define the degrees of theft as sets in their own right:

$$M = \{x: x \text{ is a misdemeanor theft}\}$$

$$N = \{x: x \text{ is a third-degree felony theft}\}$$

$$P = \{x: x \text{ is a petty misdemeanor theft}\}$$

The relationships among all the sets we have defined, both between each other and in relation to the universal set U , are best modeled with the set-theoretic relation of inclusion. Based on the text of the Model Penal Code, *any* of the material facts going to degree is sufficient to determine the degree of the offense; this *disjunctive* relationship means that we can use the previously discussed operations of union and subtraction to model how each of the material facts going to degree relate to classification of a particular case within a particular degree of offense:

$$N = E \cup F \cup G$$

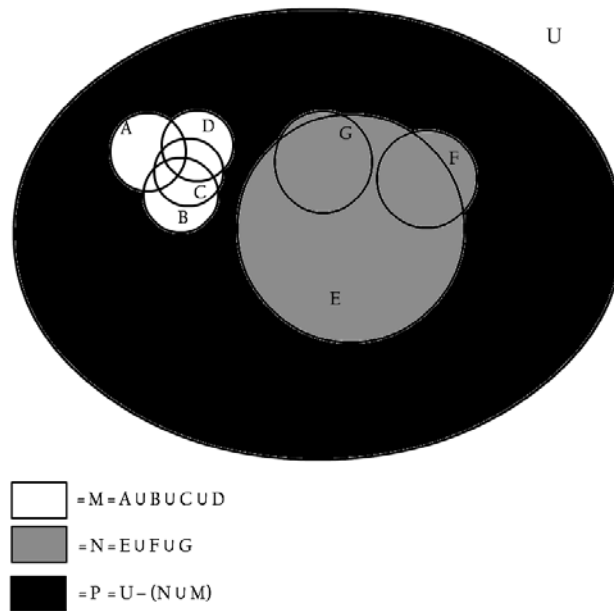
$$M = A \cup B \cup C \cup D$$

$$P = U - (N \cup M)$$

And because each of these three sets is composed of the union of other sets, we can understand each of the sets defined by the existence of a material fact going to degree to be a subset of a set defining the degree of crime committed by the defendant, which in turn is a subset of the set of the universe of all theft offenses.

Graphically, we can represent this universe of theft offenses as follows:

Figure 9



One feature of our graphical representation bears mention, and perhaps qualification. Figure 9 assumes that there is no intersection between the sets defining material facts going to degree, and therefore no intersection between the sets defining degrees of theft offenses. For example, this graphic model assumes that there is no case that falls within both set *F* (theft of a firearm) and set *A* (theft of property with value between \$50 and \$500). This assumption is not necessarily correct—indeed, it almost certainly is not. We could revise our model of this universe of discourse to account for the possibility that these two sets would intersect—that firearms worth less than \$500, or even less than \$50, could be stolen. But then we would require some basis for determining which *degree* the offense of stealing such a firearm would fall into—whether, for example, the theft of a firearm worth \$450 is a third-degree felony, or a misdemeanor, or both, or neither. Thus, the fact that our parent sets distinguishing one legal category from another—in this case, degrees of theft offenses—are composed of subsets that may intersect—in this case, sets defined by the existence of material facts going to degree—presents the potential for logical inconsistency in our model.

The Model Penal Code expressly contemplates the possibility that a single set of facts (or state of the world) might give rise to multiple offenses, or multiple degrees of an offense, simultaneously. It further seems to provide that a prosecutor has discretion to charge all crimes supported by the facts, even if this results in charging multiple degrees of the same offense, while the jury may only convict a defendant of one grade of offense “included” within the charged offense⁷⁰—a construction that will be recognized as having affinities with the set-theoretic notion of inclusion. Thus, in the case of lesser included offenses, the Model Penal Code provides explicit strategies for navigating the logical inconsistency that might result from the intersection of sets defining different grades of property crime, and treats these strategies *as legal rules in their own right*.

The important point for the present is not that the Model Penal Code solution is the correct strategy for dealing with the potential for logical inconsistency in a set-theoretic model of an area of legal doctrine, but that *some such strategy is required*. When two sets that correspond to categories defined by legal rules intersect, and those rules dictate mutually incompatible outcomes of a particular case, *some* additional content is required in order to avoid logical inconsistency. Again, we will more fully explore the possible strategies for avoiding such inconsistency in the following Part. Before doing so, we continue to fill out our basic model of legal reasoning and analysis by looking to the last of our set-theoretic operations: intersection.

3. Intersection: Legal Elements and Multi-Pronged Rules

Intersection is the set-theoretic function most intimately tied to another important type of legal structure: the concept of “elements” of a claim, defense, or legal test. Consider the elements of the tort claim analogous to the crime of theft discussed above: conversion. One of the pithier recitations of the elements of this claim states them as follows: “(1) the plaintiff has a property interest and (2) the defendant deprives the plaintiff of that interest.”⁷¹ While more complex formulations of the

⁷⁰ MODEL PENAL CODE § 1.07 (AM. LAW INST. 1962).

⁷¹ *Lassen v. First Bank Eden Prairie*, 514 N.W.2d 831, 838 (Minn. Ct. App. 1994).

conversion tort abound,⁷² this one will do to illustrate the role of intersection in our model.

When we are dealing with legal tests framed in terms of “elements,” we are typically invoking notions of both logical sufficiency and logical necessity. When we say that certain elements are *necessary* to a legal claim, we mean that the absence of *any* of those elements would be fatal to that claim. When we say that certain elements are *sufficient* to establish a legal claim, we mean that if each and every one of those elements is proven, then the absence of any *additional* fact will not cause the claim to fail. Thus, when we say that our two criteria are elements of a cause of action for conversion, what we mean is that it is *necessary* for both of them to be present in a particular state of the world subject to a legal dispute in order for us to conclude that the defendant has committed that tort,⁷³ and that the presence of both of these elements is a *sufficient* basis to conclude that the defendant committed that tort.⁷⁴ In

⁷² See, e.g., *G.S. Rasmussen & Associates, Inc. v. Kalitta Flying Serv., Inc.*, 958 F.2d 896, 906 (9th Cir. 1992) (“In California, conversion has three elements: ownership or right to possession of property, wrongful disposition of the property right and damages.”); *Cirrincione v. Johnson*, 703 N.E.2d 67, 70 (Ill. 1998) (“To prove conversion, a plaintiff must establish that (1) he has a right to the property; (2) he has an absolute and unconditional right to the immediate possession of the property; (3) he made a demand for possession; and (4) the defendant wrongfully and without authorization assumed control, dominion, or ownership over the property.”); *Vigilant Ins. Co. of Am. v. Hous. Auth. of City of El Paso*, 660 N.E.2d 1121, 1126 (N.Y. 1995) (“Conversion is the unauthorized assumption and exercise of the right of ownership over goods belonging to another to the exclusion of the owner’s rights.”) (internal quotation marks omitted).

⁷³ Thus, in jurisdictions where conversion of intangible property is actionable and where damage is an element of the cause of action, a plaintiff who shows ownership of the intangible property and the defendant’s taking of a copy of that property nevertheless cannot establish a conversion claim if the defendant destroys his copy of the plaintiff’s property prior to making any use of it. See, e.g., *News Am. Mktg. In-Store, Inc. v. Marquis*, 862 A.2d 837, 848 (Conn. App. Ct. 2004), *aff’d*, 885 A.2d 758 (Conn. 2005).

⁷⁴ Thus, the fact that a conversion defendant lacked any bad-faith intent—or *mens rea*—is generally no barrier to a conversion claim. *Deming v. Nationwide Mut. Ins. Co.*, 905 A.2d 623, 639–40 (Conn. 2006) (“[S]tatutory theft requires an intent to deprive another of his property . . . [t]herefore, statutory theft requires a plaintiff to prove the additional element of intent over and above what he or she must demonstrate to prove conversion.”); *Ahles v. Aztec Enters., Inc.*, 502 N.Y.S.2d 821, 822 (N.Y. App. Div. 1986) (“Intent to possess another’s property is not an essential element of conversion.”); DAN B. DOBBS, PAUL T. HAYDEN & ELLEN M. BUBLICK, *THE LAW OF TORTS* § 62 (2d ed. 2018 update) (“The intent required to show conversion is exactly analogous to the intent required to prove a trespass to land. In neither

other words, the tort is established by a *conjunction* of these elements, and as discussed above, the logical relation of conjunction is expressed in set theory through the concept of *intersection*.⁷⁵

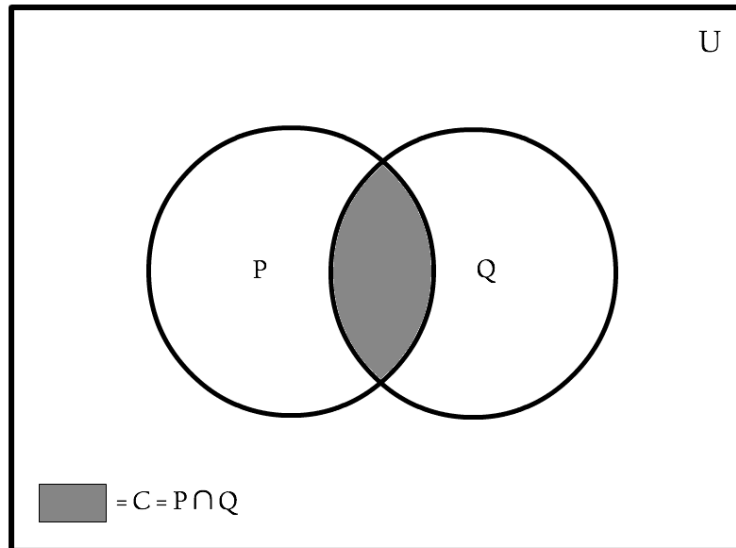
Once again, we can develop this concept both formally and graphically. Consider a universe of discourse consisting of states of the world in which someone has a legally cognizable property interest in a particular object of interest. Let us further define two sets within this universe of discourse by reference to the parties to a legal dispute. Let P be the set of all cases in which the plaintiff has a property interest in the object of interest, and let Q be the set of all cases in which the defendant has prevented others from exercising control over the object of interest. We can understand viable conversion claims as those that fall within the intersection of these two sets:

$$C = P \cap Q$$

Graphically, we can represent this simplified model of conversion as follows:

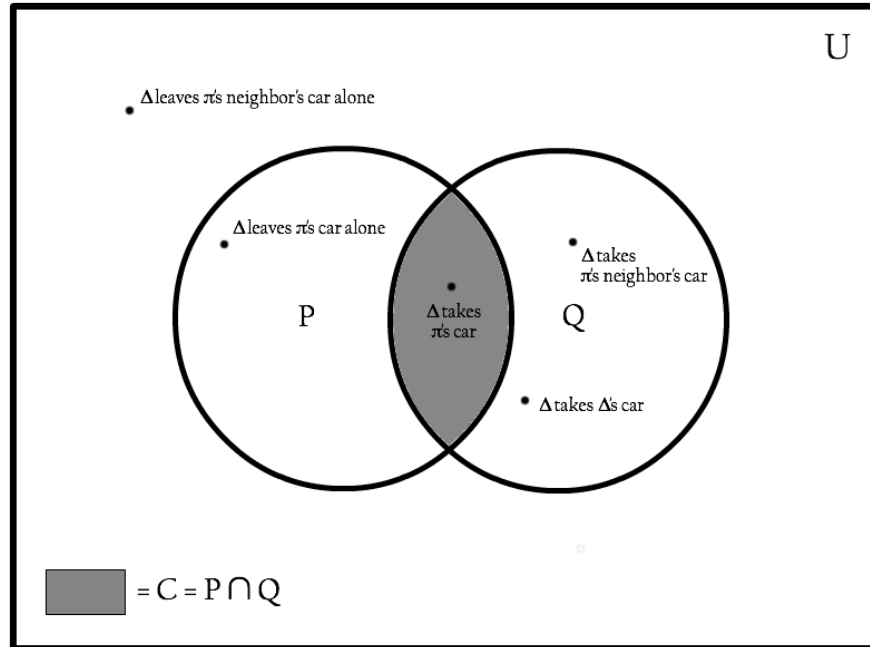
case is the defendant's bad motive or good faith ordinarily relevant except on the question of punitive damages.").

⁷⁵ There are computational theories of legal reasoning that take a similar approach to more complex modeling problems by conceiving of cases as collections of "factors"—facts relevant to a rule directing a disposition. See generally KEVIN D. ASHLEY, MODELING LEGAL ARGUMENTS: REASONING WITH CASES AND HYPOTHETICALS (1991); John F. Horty, *Rules and Reasons in the Theory of Precedent*, 17 LEGAL THEORY 1 (2011).

Figure 10

Thus, property interests that the plaintiff has in an object of interest, but which the defendant has not prevented the plaintiff from exercising control over, do not give rise to a conversion claim: The fact that you stole my car doesn't mean I can sue you for conversion of the bicycle I still have. Likewise, objects that the defendant prevents others from enjoying will not ground a conversion claim by plaintiffs who have no property interest in those objects: You can't sue me for conversion for driving my own car, or even your neighbor's car, no matter how much you might disapprove of my conduct, or wish the car were yours. Again, we can present various factual scenarios graphically as part of our set-theoretic model of the relevant doctrine:

Figure 11



As should be apparent, the predicates of these two sets are not especially informative. What it means for a plaintiff to have a property interest in a thing, and what it means for a defendant to prevent someone from exercising control over a thing, are not questions that our model can necessarily answer. And while this problem is not equally vexing for all predicates that might serve as intensional definitions of sets of legal interest—for example, “ x is a natural person more than 18 years of age”—it is quite a common concern in legal analysis. To solve such a problem, the common law often turns from *intensional* to *extensional* definition—a mode of reasoning that will complete our basic model of set theory in legal analysis.

C. Common-Law Rule-Building: Extensional Definition

So far, our set-theoretic model of legal analysis has relied on intensionally defined sets. A legal authority (a statute or binding precedent) announces a rule that purports to define a category; subsequent cases are then evaluated for membership in that category. But in common-law systems, cases are not decided purely by deductive

reasoning from abstract rules of law encapsulated in legal authorities. They are also decided inductively, based on extrapolation of rules from the facts and outcomes of analogous precedents. In terms of legal theory, this is the interpretivist exercise of “discover[ing] principles that *fit*, not only the particular precedent to which some litigant directs his attention, but all other judicial decisions within [the court’s] general jurisdiction.”⁷⁶ In our set-theoretic model, we can understand it as an exercise in *extensional* set definition.⁷⁷ Importantly, this exercise interacts with the previously discussed exercise of determining whether a particular state of the world falls within an *intensional* definition. Moreover, this interaction between intensional and extensional set definition is *recursive*: intensionally defined rules guide the determination of cases, and the determination of cases goes on to inform the construction (and reformation) of intensional definitions.

This recursive dynamic is nicely illustrated in the Ninth Circuit opinion in *Kremen v. Cohen*.⁷⁸ Fortuitously, this case about conversion of a domain name also concerns an issue left open in the previous Section: how we might give content to a vague or ambiguous intensional definition such as “a property right of the plaintiff” in a marginal case. In *Kremen*, the dispute turned in part on a discrete legal issue: whether an internet domain name is “property” for purposes of the California law of conversion. To answer this question, Judge Kozinski,⁷⁹ writing for the panel, considered two available *intensional* predicates for the relevant definition of “property” as applied to intangibles. One earlier Ninth Circuit case defined “property” as including “every intangible benefit and prerogative susceptible of possession or disposition.”⁸⁰ In contrast, the Restatement (Second) of Torts states that intangible rights

⁷⁶ Ronald Dworkin, *Hard Cases*, 88 HARV. L. REV. 1057, 1094 (1975) (emphasis added).

⁷⁷ Schauer’s discussion of the common law, for example, includes a discussion of the extensional definition of rules based on the decisions of cases. SCHAUER, *supra* note 9, at 183–85.

⁷⁸ 337 F.3d 1024 (9th Cir. 2003).

⁷⁹ In 2009, Judge Kozinski was admonished by a judicial disciplinary panel for maintaining a publicly accessible server that included sexually explicit material. In 2017, he abruptly resigned after being accused of sexual misconduct by numerous women, including former law clerks.

⁸⁰ *Kremen*, 337 F.3d at 1030 (quoting *Downing v. Mun. Court*, 198 P.2d 923, 926 (Cal. Ct. App. 1948)).

could be property subject to conversion “[w]here there is conversion of a document in which intangible rights are merged” or where the defendant “effectively prevents the exercise of intangible rights of the kind customarily merged in a document,” and further explains that “[a]n intangible is merged in a document when . . . the right to the immediate possession of a chattel . . . is represented by [the] document, or when an intangible obligation [is] represented by [the] document, which is regarded as equivalent to the obligation.”⁸¹ The choice of either of these definitions could be outcome determinative. After all, control over a domain name—which is no more or less than the system of distributed computer data records and architectural conventions that allows a particular string of text to uniquely identify a computer on the internet⁸²—might be an “intangible benefit susceptible of . . . disposition,” but nevertheless might not be “merged in a document.”

We can model this choice using the set-theoretic tools developed above. We can first illustrate the relevant universe of discourse as divided into our two complementary sets, π and Δ , corresponding to a judgment for the plaintiff and a judgment for the defendant, respectively. Given the *Kremen* court’s determination that a conversion claim under California law requires the plaintiff to show “ownership or right to possession of property, wrongful disposition of the property right and damages,”⁸³ we can define the set corresponding to a judgment for the plaintiff as the intersection of three sets corresponding to these three elements. Taken together, we have the following formal definitions:

⁸¹ *Id.* at 1031 (emphasis omitted) (internal quotation marks omitted) (quoting RESTATEMENT (SECOND) OF TORTS § 242 & cmt. a (AM. LAW INST. 1965)).

⁸² See *id.* at 1033–35; see also INTERNET ENGINEERING TASK FORCE, RFC-1035: DOMAIN NAMES – IMPLEMENTATION AND SPECIFICATION (Nov. 1987), <https://tools.ietf.org/html/rfc1035> [<https://perma.cc/KLH8-T393>].

⁸³ *Kremen*, 337 F.3d at 1029 (quoting *G.S. Rasmussen & Assocs., Inc. v. Kalitta Flying Serv., Inc.*, 958 F.2d 896, 906 (9th Cir. 1992)).

$f(x) = x$ is a property right owned by the plaintiff

$$P = \{x: f(x)\}$$

$Q = \{x: x \text{ is a property right wrongfully disposed of by the defendant}\}^{84}$

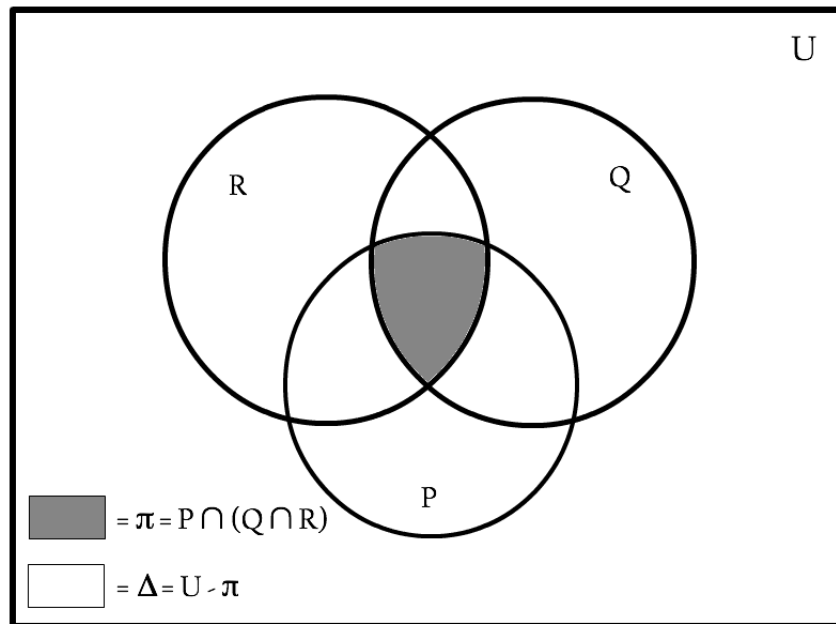
$R = \{x: x \text{ is an act of the defendant that causes damage to the plaintiff}\}$

$$\pi = P \cap (Q \cap R)$$

$$\Delta = (U - \pi) = \pi'$$

We can also represent this universe of discourse graphically as follows:

Figure 12



The main dispute in *Kremen* was over the size (or scope) of P . As the discussion above indicates, at least two alternative understandings of the predicate for P were available. We can amend our model to account for these alternative predicates:

⁸⁴ Obviously, we could also define Q as the intersection of two other sets: property rights disposed of by the defendant and property rights the defendant has no right to dispose of. But doing so adds no clarity to the analysis in this case. It might be an important distinction, however, in cases where the defendant asserts such a right (for example, if the defendant asserts permission, ownership, necessity, or some other form of justification or excuse for the disposition).

$g(x) = x$ is “any intangible benefit and prerogative susceptible of possession or disposition” owned by the plaintiff

$$P_1 = \{x: g(x)\}$$

$h(x) = x$ is a right to tangible property owned by the plaintiff

$i(x) = x$ is an intangible right owned by the plaintiff which is “represented by a document”⁸⁵

$$P_2 = \{x: (h(x) \vee i(x))\}$$

$$\pi_1 = (P_1 \cap Q) \cap R$$

$$\Delta_1 = (U - \pi_1) = \pi_1'$$

$$\pi_2 = (P_2 \cap Q) \cap R$$

$$\Delta_2 = (U - \pi_2) = \pi_2'$$

Judge Kozinski seemed to be operating under the premise that the facts of *Kremen* would satisfy the Ninth Circuit’s definition of property, but not the Restatement’s definition.⁸⁶ This is a consequential—and perhaps contestable—premise, which we will not examine here, saving such evaluation of membership relations for further investigation in Part III. For now, Judge Kozinski’s premise is mainly relevant in that it provides some additional propositions for our model:

$$k = \textit{Kremen v. Cohen}$$

$$k \in P_1$$

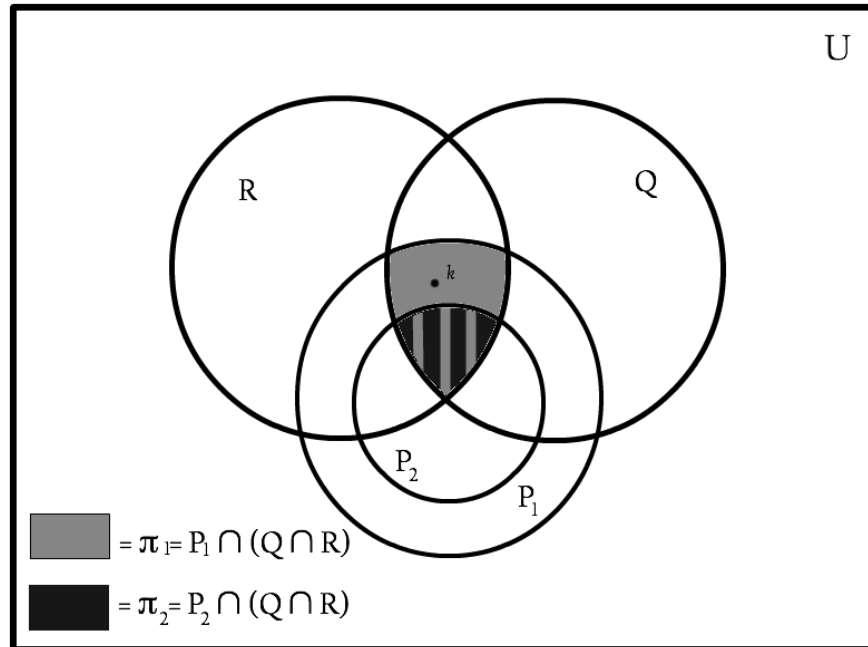
$$k \notin P_2$$

Given these revised and additional definitions, we can revise our graphic representation of the doctrine of conversion at issue in the case:

⁸⁵ This predicate is the most appropriate for the analysis that immediately follows, but we could obviously draw further distinctions concerning the alleged wrongful act of the defendant, such as whether the defendant interfered with the document representing the intangible right or merely with the right itself.

⁸⁶ *Kremen*, 337 F.3d at 1030–31.

Figure 13



If—as our model suggests—the outcome of *Kremen* depended on whether P_1 or P_2 is the appropriate category of “property” for purposes of establishing a California conversion claim, Judge Kozinski’s task was to decide which category was required by California law. To do so, he engaged in the classic common-law exercise of reviewing analogous precedent. Three cases decided by California courts particularly drove the analysis. *Payne v. Elliot*,⁸⁷ the only California Supreme Court case on the issue, stated that the action for trover (the common-law predecessor of conversion) was “a remedy for the conversion of every species of personal property,” and therefore concluded that shares in a corporation could be converted even without conversion of the actual stock certificates themselves.⁸⁸ *Olschewski v. Hudson*, in contrast, stated that “the proceeding in conversion was not intended to reach so intangible, uncertain, and indefinite a property right” as the interest the plaintiff in that case was claiming as property: a laundry route.⁸⁹ In

⁸⁷ 54 Cal. 339 (1880).

⁸⁸ *Id.* at 341.

⁸⁹ *Olschewski v. Hudson*, 262 P. 43, 45–46 (Cal. Dist. Ct. App. 1927).

doing so, *Olschewski* remarked that *Payne*'s categorical statement about the scope of conversion was "too broad" and "unnecessary to the determination of the issue in that case," which could have been resolved on grounds that "[s]hares of stock are represented by certificates which are evidence of a definite interest in the assets of a company" and are therefore "tangible."⁹⁰ *Olschewski* thus relied on a rule quite similar to that of the Restatement. But a later case, *Palm Springs-La Quinta Development Co. v. Kieberk Corp.*,⁹¹ distinguished *Olschewski* to uphold a conversion claim against a defendant who wrongfully took possession of a large set of index cards recording customer information, and subsequently lost or destroyed a large number of those cards. The *Palm Springs* court's justification for departing from *Olschewski* seemed to rest on the fact that the physical index cards themselves were tangible in a way that the goodwill and customer lists of a laundry service were not.⁹²

Based on these and other cases, Judge Kozinski drew distinctions between the *intensional* definitions provided by the language of the reviewed opinions, and the *extensional* definitions that could be inferred by grouping cases according to their outcomes. With respect to *Payne*, for example, he noted that "[w]hile *Payne*'s outcome might be reconcilable with the *Restatement*, its rationale certainly is not: It recognized conversion of shares, not because they are customarily represented by share certificates, but because they are a species of personal property and, perforce, protected."⁹³ *Olschewski* presented the

⁹⁰ *Id.* at 46. Another laundry-route case relied on *Olschewski* to arrive at the same conclusion. *Adkins v. Model Laundry Co.*, 268 P. 939, 942 (Cal. Ct. App. 1928) ("[T]here is no such property right in the intangible interest of an exclusive privilege to collect laundry or sell newspapers in a specific district, which will authorize damages in a suit at law for conversion or trover.").

⁹¹ 115 P.2d 548 (Cal. Dist. Ct. App. 1941).

⁹² *Id.* at 552 ("The defendant in [*Olschewski*] was not charged with converting an index list of customers. He was charged with selling the good-will of a business in a specified district. This court merely held that was an intangible property right which is not susceptible of conversion. In the present case the appellants were charged with damages for destroying and appropriating tangible personal property consisting of a cabinet of lead cards containing the names and valuable information regarding prospective and actual purchasers of real property, contrary to the express terms of a written contract. There is a clear distinction between these cases.").

⁹³ *Kremen v. Cohen*, 337 F.3d 1024, 1031 (9th Cir. 2003).

opposite problem. The language of that opinion raised “a plausible argument that California follows the *Restatement*.”⁹⁴ But this argument was undermined by comparison of the result in *Olschewski* with the result in *Palm Springs*. As Judge Kozinski put it:

Palm Springs and *Olschewski* are reconcilable on their facts—the former involved conversion of the document itself while the latter did not. But this distinction can’t be squared with the *Restatement*. The plaintiff in *Palm Springs* recovered damages for the value of his intangibles. But if those intangibles were merged in the index cards . . . the plaintiffs in *Olschewski* . . . should have recovered [as well] . . . [L]aundry routes surely are customarily written down somewhere.⁹⁵

In attempting to reconcile these three cases, Judge Kozinski placed the greatest weight not on the rationales (or rules) announced in the judicial opinions he reviewed, but on the correlation of their underlying facts with their outcomes. Specifically, he tried to determine whether that correlation was captured by one of the available intensional definitions under consideration. If *Olschewski* held for the defendant on a rationale that was inconsistent with the outcome of *Palm Springs*, and if *Payne* held for the plaintiff on a rationale that was inconsistent with the outcome of *Olschewski*, both rationales should be disregarded in favor of a *new* rationale that can explain the outcomes in *all three* cases by reference to their underlying facts:

To the extent *Olschewski* endorses the strict merger rule, it is against the weight of authority. . . . Were it necessary to settle the issue once and for all, we would toe the line of *Payne* and hold that conversion is “a remedy for the conversion of every species of personal property.” But we need not do so to resolve this case. Assuming *arguendo* that California retains some vestigial merger requirement, it is clearly minimal, and at most requires only *some* connection to a document or tangible object—not representation of the owner’s intangible interest in the strict *Restatement* sense.⁹⁶

⁹⁴ *Id.* at 1032.

⁹⁵ *Id.*

⁹⁶ *Id.* at 1033 (internal citation omitted).

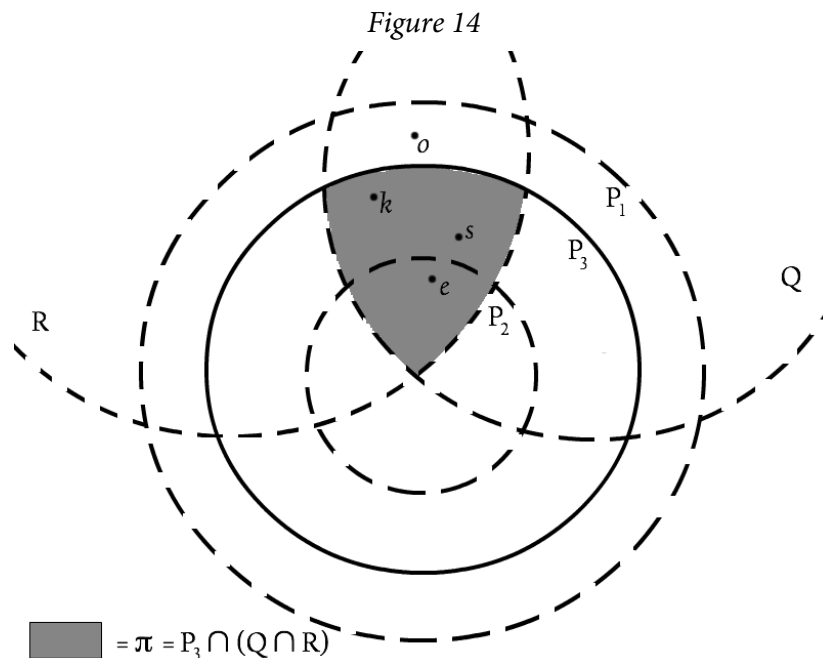
Having thus (re)defined the relevant scope of a “property right,” Judge Kozinski went on to conclude that a domain name is sufficiently connected to a “document”—specifically, information stored in a collection of electronic records comprising the Domain Name System—to satisfy his definition and thus make out a case for conversion.⁹⁷

We are now at last in a position to model the foregoing exercise in extensional set building both formally and graphically. We have three new cases to categorize, and must also formulate a new predicate for the set corresponding to the appropriate definition of “property.” Let us call that set P_3 . We thus have the following formal definitions:

$$\begin{aligned} e &= \textit{Payne v. Eliot} \\ o &= \textit{Olschewski v. Hudson} \\ s &= \textit{Palm Springs Development Co. v. Kieberk Corp.} \\ P_3 &= \{x: (h(x) \vee j(x))\} \\ e &\in P_3 \\ s &\in P_3 \\ k &\in P_3 \\ o &\notin P_3 \end{aligned}$$

Our task now is to formulate the as-yet-unspecified predicate $j(x)$, which defines the type of *intangible* right that is “property” under California conversion law. Because we are proceeding by extensional definition, it may be most helpful to turn here to our graphical model, focusing in on the area of interest. If we map our cases according to the relationship between the intangible right at issue and some document (with placement closer to the center of P implying a closer relationship to a traditional document), we might arrive at something like this:

⁹⁷ *Id.* at 1033–34.



The task, then, is to identify a predicate $j(x)$ that is satisfied by e , k , and s , but is *not* satisfied by o . Judge Kozinski's solution to this problem can be formally expressed as follows:

$j(x) = x$ is an intangible right having some minimal connection to a document or tangible object

This type of exercise—organizing cases according to their outcomes and identifying some common feature that is present in all cases with one outcome but absent from all cases with the opposing outcome—is the hallmark of common-law reasoning by analogy and distinction. It should be familiar to most lawyers, and is among the key skills to be mastered by law students. But conceiving of this classic exercise in set-theoretic terms highlights some important points.

First: The relationship between rules and cases is both *dynamic* and *recursive*. Announced intensional rules inform the outcomes of individual cases, but extensional definitions based on the outcomes of those individual cases may then in turn qualify, modify, and potentially even overturn those intensional rules, and so on and so forth. This

recursive process generates constant, dynamic refinement and reformulation of the universe of discourse of legal doctrines.⁹⁸

Second: Judge Kozinski gives extensional definition precedence over intensional definition. Indeed, his effort to reconcile *Payne* with inferior California court opinions suggests that this precedence may be *more* important in his view than the transmission of intensional rules via the hierarchies of appellate authority. We can understand this preference for extensional definition as a reflection of Dworkin's observation about the notion of "fit": that a judge "does not satisfy his duty to show that his decision is consistent with established principles, and therefore fair, if the principles he cites as established are themselves inconsistent with other decisions that his court also proposes to uphold."⁹⁹ As in the context of lesser included offenses, a conflict between two applicable rules requires some strategy for resolution. The interpretivist concepts of fit and justification appear to constitute another such strategy: one that purports to resolve contradictions between intensional definitions from the text of judicial opinions and extensional definitions derived from the juxtaposition of multiple prior judicial decisions in favor of the latter. Again, the point is not that Judge Kozinski's (or Dworkin's) strategy is the right one, but that some such strategy is required.

Third: The prioritization of one rule over another can be implemented in a number of formally distinct ways, with differing implications for the legal universe of discourse in which the conflict of rules arises. As Judge Kozinski recognized, he could simply have followed the intensional definition of "property" extracted from the *Payne* opinion, and evaluated Kremen's claimed property interest against that definition. This would have entailed a conclusion that *Olschewski* was wrongly decided, and reordered the membership relations between the discussed cases and the intensionally-defined set containing them. Instead of doing so, Judge Kozinski altered the inclusion relationships between the sets defined by intensional predicates from earlier judicial opinions and the sets defining the outcomes of conversion cases, adding a new intensionally-defined set of

⁹⁸ Cf. Frederick Schauer, *Precedent*, 39 STAN. L. REV. 571, 572-73 (1987) (noting that precedent is both "forward-looking" and "backward-looking").

⁹⁹ Dworkin, *supra* note 76, at 1094.

his own to the universe of discourse. Rather than altering the *membership* relations between the precedents and rule-based sets, Judge Kozinski altered the *inclusion* relationships between rule-based sets and the outcomes of adjudication. These strategies are distinct, not in their result for the particular case (Kremen would have prevailed either way), but as a matter of logical *form*. Even so, as precedents in their own right, opinions implementing such strategies make different claims on the *future* legal universe of discourse.

This type of formal difference in the behavior of legal actors becomes more complex when we consider that not all legal authorities lend themselves to this type of extensional set-building. In particular, regulations, statutes, and constitutions cannot be analyzed extensionally, because they consist solely of intensional definitions. They have no facts, no outcomes, to feed into the recursive process of extensional construction. Still, when a conflict arises among such intensional definitions—whether derived from non-caselaw authorities or from the language of judicial opinions—there may be other ways of avoiding a logically inconsistent set-theoretic model of the universe of discourse subject to those authorities. Mapping out those strategies—and their implications—is the project of the rest of this Article.

III. FORMAL STRATEGIES FOR AVOIDING CONTRADICTION

The set-theoretic tools developed in the previous Part may be helpful in formalizing ordinary legal reasoning, but if that were all they achieved they would mainly be of interest to only a small set of theorists, educators, and (perhaps) law students. After all, most lawyers and judges get by perfectly well in their professional lives without conceiving of their arguments and analyses in terms of predicate logic or Venn diagrams. In this Part, however, I will argue that the tools developed in the previous Part are necessary to adequately understand an important feature of legal practice that is intimately connected to the process of common-law doctrinal development and change.

Specifically, I propose that legal actors are free to respond in particular cases to the constraints imposed by conflicting applicable legal authorities by resorting to at least three formally distinct types of arguments: (1) arguments about the relation of those authorities to the facts of particular cases; (2) arguments about the relation of those

authorities to one another; or (3) arguments about the relation of those authorities to adjudicative outcomes. The distinctions among these types of arguments map to discrete aspects of the set-theoretic model developed in the previous part, but *not necessarily* to discrete features of the natural language in which legal authorities are expressed or to the substance of the rules embodied in those authorities. Thus, understanding the set-theoretic model of legal reasoning allows for deeper appreciation of the choices made by legal actors in advocacy and adjudication, and raises theoretical issues separate from typical jurisprudential concerns founded on the ambiguity (or “open texture”¹⁰⁰) of language. This Part focuses on the first of these issues—understanding how strategies for resolving doctrinal conflicts differ, and how those differences relate to changes in doctrine over time. The next Part will address the second issue: how a formal understanding of the diverse strategies for negotiating doctrinal conflict relates to legal theoretical debates—particularly those focusing on the determinacy (or lack thereof) of law.

A. *Competing Predicates and Logical Contradictions: Outlining the Problem*

Distinctions among the logical forms of strategies for resolving doctrinal conflicts are best illustrated with a simple model of contradiction and constraint: the apparent applicability of two legal rules requiring opposite outcomes to the facts of a single case. We first addressed this possibility explicitly when discussing lesser included criminal offenses;¹⁰¹ we will see shortly that our discussion of subleases and assignments presented a similar problem.¹⁰² When presented with such a conflict of authority, a legal actor may struggle to maintain the logical consistency of a body of doctrine. The resolution of that conflict requires legal actors to manipulate the *logical structure* of the legal universe of discourse, and there are multiple ways of doing so. In short,

¹⁰⁰ HART, *supra* note 8, at 127–28; see generally Frederick Schauer, *On the Open Texture of Law*, 87 GRAZER PHILOSOPHISCHE STUDIEN 197 (2013).

¹⁰¹ *Supra* Section II.B.2.

¹⁰² *Supra* Section II.B.1.

legal actors have significant discretion in the *formal* strategies they use in arguing or adjudicating cases.

To see how the conflict of legal authorities generates multiple possible *formally distinct* resolution strategies, let us consider some examples from intellectual property law. Intellectual property lawyers and scholars often group legal rules in the field into four distinct groups: validity, infringement, defenses, and remedies.¹⁰³ Questions of validity—whether someone has a protectable intellectual property right or not—particularly lend themselves to set-theoretical analysis.

Let us suppose that when asking whether an individual has a valid intellectual property right or not (which we may for present purposes consider a binary proposition¹⁰⁴), we are asking whether the intangible asset they claim satisfies certain predicates for set membership. Specifically (following the analysis of the previous Part), we are asking whether the intangible asset is a member of a set defined as the intersection of other sets, which are in turn defined by the applicable legal rules concerning validity. For copyrights, we ask whether the claimed right is protectable subject matter (i.e., a work of authorship),¹⁰⁵ whether it is original,¹⁰⁶ and whether it is fixed in a tangible medium of expression¹⁰⁷—an intersection of three sets. For trademarks, we ask

¹⁰³ Mark A. Lemley & Mark P. McKenna, *Scope*, 57 WM. & MARY L. REV. 2197, 2200 & n.1 (2016) (“Intellectual property (IP) law doctrines fall into three basic categories: validity, infringement, and defenses. Virtually every significant legal doctrine in IP is either about whether the plaintiff has a valid IP right that the law will recognize (validity); whether what the defendant did violates that right (infringement); or whether the defendant is somehow privileged to violate that right (defenses). . . . If the IP owner prevails, there are also issues about the remedy awarded.”).

¹⁰⁴ Lemley and McKenna argue that many intellectual property law cases turn not on the validity of an intellectual property right but on its scope—a non-binary proposition. *See generally id.* While this is true, it is always possible to reformulate the non-binary question of a right’s scope as a binary question: i.e., “does the plaintiff have a right *of this scope* or not”? *See supra* notes 1–4 and accompanying text. Importantly, this is a formal move, not a substantive one—it reframes the question in such a way as to allow clearer modeling of the applicable analysis according to the tools used in this Article, but still requires some substantive content as to the appropriate scope of intellectual property rights, and moreover requires some external reason for selecting a right of the posited scope as the appropriate subject for analysis.

¹⁰⁵ 17 U.S.C. § 102 (1990).

¹⁰⁶ *Feist Pubs., Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345–47 (1991).

¹⁰⁷ 17 U.S.C. § 101 (2010); *Cartoon Network LP v. CSC Holdings, Inc.*, 536 F.3d 121, 129 (2d Cir. 2008) (“[T]o determine whether a work is ‘fixed’ in a given medium, the statutory language

whether the claimed mark is protectable subject matter,¹⁰⁸ whether it is distinctive,¹⁰⁹ and whether the claimant has used it in commerce¹¹⁰—another intersection of three sets. For patents, we ask whether the claimed invention is patentable subject matter, novel, useful, non-obvious, and has been adequately described in an enabling disclosure—an intersection of five sets.¹¹¹ To reiterate: This last example envisions that inventions that can be categorized as to their membership in the set defined as the intersection of five other sets, themselves defined by the substantive legal criteria of patentability. Inventions within that area of intersection are entitled to patent protection—we may say they are elements of the set of patentable inventions—while those that are outside that set are unpatentable (or invalid if a patent has erroneously issued for them).

Moreover, each of the five criteria for patentability (whose intersection determines a patent's validity) may be further defined via set-theoretic operations. Consider patentable subject matter. Section 101 of the Patent Act defines patentable subject matter as any “process, machine, manufacture, or composition of matter.”¹¹² We can express this rule formally using the tools developed in the previous Part. First, we must establish some definitions:

directs us to ask not only 1) whether a work is ‘embodied’ in that medium, but also 2) whether it is embodied in the medium ‘for a period of more than transitory duration.’”).

¹⁰⁸ *Qualitex Co. v. Jacobson Prods. Co.*, 514 U.S. 159, 162 (1995) (“[T]he Lanham Act . . . says that trademarks ‘includ[e] any word, name, symbol, or device, or any combination thereof.’ [15 U.S.C.] § 1127. Since human beings might use as a ‘symbol’ or ‘device’ almost anything at all that is capable of carrying meaning, this language, read literally, is not restrictive.”).

¹⁰⁹ *Two Pesos, Inc. v. Taco Cabana, Inc.*, 505 U.S. 763, 768 (1992) (holding that to be protected, a trademark must be “capable of distinguishing the applicant’s goods from those of others”).

¹¹⁰ 15 U.S.C. § 1127 (2006) (“The term “trademark” includes any word, name, symbol, or device, or any combination thereof . . . *used* by a person . . . to identify and distinguish his or her goods.” (emphasis added)); *Hydro-Dynamics, Inc. v. George Putnam & Co.*, 811 F.2d 1470, 1473 (Fed. Cir. 1987) (“The requirements of both adoption and use devolve from the common law; trademark rights in the United States are acquired by such adoption and use, not by registration.”).

¹¹¹ 35 U.S.C.A §§ 101–103; 112 (West 2015).

¹¹² 35 U.S.C.A § 101 (West 2015). To simplify this illustrative example we will omit the patentability of improvements to existing inventions.

$$\begin{aligned}
 P(x) &= x \text{ is a process} \\
 Q(x) &= x \text{ is a machine} \\
 R(x) &= x \text{ is a manufacture} \\
 S(x) &= x \text{ is a composition of matter} \\
 A &= \{x : P(x)\} \\
 B &= \{x : Q(x)\} \\
 C &= \{x : R(x)\} \\
 D &= \{x : S(x)\}
 \end{aligned}$$

Now let us imagine a set E representing the set of all inventions that qualify as patentable subject matter. Given our prior definitions and the text of Section 101 we can construct E as a *union* of several other sets:

$$E = A \cup B \cup C \cup D$$

The statutory definition of patentable subject matter is *inclusionary*, i.e., it tells us what *is* patentable subject matter. One might therefore think that everything outside this definition is *not* patentable subject matter; i.e., that the complement E' would be a complete and adequate formal description of subject matter that is unpatentable under Section 101. But it isn't. And this is because the universe of discourse regarding patentable subject matter includes more legal sets than those encompassed by the statutory definition of Section 101.

As Justice Breyer noted in *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*,¹¹³ "The [Supreme] Court has long held that [Section 101] contains an important implicit exception. '[L]aws of nature, natural phenomena, and abstract ideas' are not patentable."¹¹⁴ This is, obviously, an *exclusionary* definition: It tells us what is *not* patentable subject matter. And we might ask whether this definition is equivalent to our previously derived description of unpatentable subject matter, E' .

All scholars and practitioners of patent law know that these two definitions are *not* equivalent, and even those unschooled in this area of doctrine will immediately understand why. The non-equivalence is signaled by Justice Breyer's use of the word "exception" in *Mayo*. But

¹¹³ 566 U.S. 66 (2012).

¹¹⁴ *Id.* at 70.

formalizing what this exception means with respect to Section 101—how the inclusionary and exclusionary definitions interact as a matter of logical form—will help us understand the value of set-theoretic analysis for managing conflicting intensional predicates. Let us construct a formal definition of *unpatentable* subject matter under the Supreme Court's *exclusionary* definition and attempt to integrate it with our formalization of the *inclusionary* definition of Section 101:

$$\begin{aligned} T(x) &= x \text{ is a law of nature} \\ V(x) &= x \text{ is a natural phenomenon} \\ W(x) &= x \text{ is an abstract idea} \\ F &= \{x : T(x)\} \\ G &= \{x : V(x)\} \\ H &= \{x : W(x)\} \end{aligned}$$

Now let us suppose some set K that will represent the set of all *unpatentable* subject matter under the Supreme Court's definition as stated in *Mayo*:

$$K = F \cup G \cup H$$

The question we are trying to answer is whether either of the following logically equivalent assertions are true:

$$\begin{aligned} E &= K' \\ K &= E' \end{aligned}$$

That is, we want to know whether the *inclusionary* definition of Section 101 and the *exclusionary* definition from Supreme Court caselaw are *complements* in the set-theoretic sense of the word.

And now the problem is clear: These two definitions are *not* complements. Moreover, the reason why they are not complements is important—it is because some of the subsets of E have non-empty intersections with some of the subsets of K . We can see such intersection in two lines of cases. The first involves cases such as *Parke-Davis & Co. v. H.K. Mulford Co.*¹¹⁵ and *Association of Molecular Pathology v. Myriad Genetics, Inc.*,¹¹⁶ in which certain isolated and purified biological products are—at least arguably—elements both of set D (compositions

¹¹⁵ 189 F. 95 (2d Cir. 1911) (Learned Hand, J.).

¹¹⁶ 569 U.S. 576 (2013).

of matter) and set G (natural phenomena). The second involves cases such as *Bilski v. Kappos*¹¹⁷ and *Alice Corp. v. CLS Bank Int'l*,¹¹⁸ in which certain computerized methods useful in the finance industry are—at least arguably—elements both of set A (processes) and set H (abstract ideas). These examples of objects that are elements of subsets of both E and K render the complementarity of E and K logically impossible: complements are necessarily disjoint sets.¹¹⁹

To illustrate the same point graphically, imagine some object n that is an element of both sets A and H —i.e., it is both a process and an abstract idea. We can represent the resulting universe of discourse as follows:

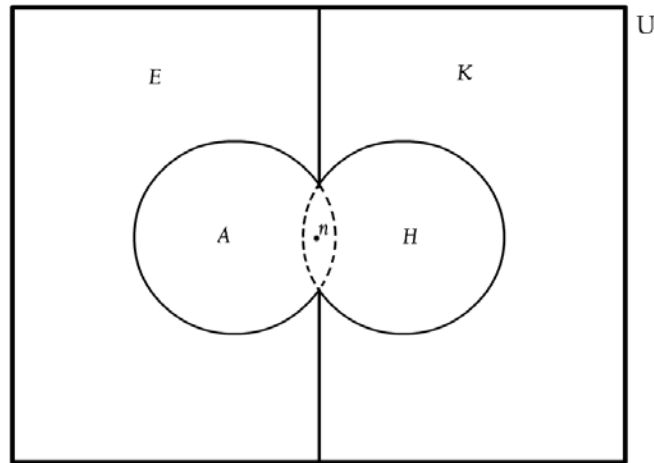
¹¹⁷ 561 U.S. 593 (2010).

¹¹⁸ 573 U.S. 208 (2014).

¹¹⁹ We can establish this by assuming that E and K are complements and then proving a contradiction based on the existence of an object that is an element of subsets of both E and K . Formally we may express our proof as follows, given a universe of discourse including the sets as defined in the text and some object n that is both a process and an abstract idea:

- | | | |
|------|--|------------------------------------|
| (1) | Let $E = K'$ | (Assumption) |
| (2) | $E = U - K$ | (Definition of Complementarity; 1) |
| (3) | $E = \{x : (x \in U) \wedge (x \notin K)\}$ | (Definition of Set Subtraction) |
| (4) | $\neg \exists (x)[(x \in E) \wedge (x \in K)]$ | (Definition of Set Membership; 3) |
| (5) | <u>$(n \in A) \wedge (n \in H)$</u> | (Given) |
| (6) | <u>$A \subseteq E$</u> | (Given) |
| (7) | <u>$H \subseteq K$</u> | (Given) |
| (8) | $(n \in E) \wedge (n \in K)$ | (Definition of Inclusion; 5, 6, 7) |
| (9) | \perp | (6, 8) |
| (10) | $\neg (E = K')$ | (1, 9) |

Figure 15



By the definition of the relation of inclusion, n will be a member of both E (which includes its subset A) and K (which includes its subset H). But if this is true, E and K cannot be complements, because complements can have no elements in common.

Thus, treating the inclusionary statutory definition and the exclusionary common-law definition of patentable subject matter as mutually consistent—i.e., as set-theoretic complements—is logically untenable. If processes are patentable, but abstract ideas are not patentable, any object that is both a process and an abstract idea yields the above-described logical contradiction—it must be both patentable and not patentable at the same time. Indeed, we can generalize this understanding. Expressed as a matter of set theory, we may say that any time two definitions for a legal set exist—one inclusionary and one exclusionary—if the definitions do not define *disjoint sets* they will generate logical contradictions. The question then arises how to avoid such contradictions. We will consider three strategies for doing so: (1) the Trump Card, (2) the No True Scotsman, and (3) the Tertium Quid.

B. *The Trump Card*

As suggested earlier, the patent lawyer has a ready answer to the accusation of logical inconsistency in patentable subject matter doctrine. This answer is implied in the recitation of the exclusionary definition from *Mayo* quoted above: that the rule of Section 101 “contains an

important implicit *exception*.”¹²⁰ That is, the Supreme Court may be understood as saying that its definition is *an exception to the otherwise applicable* statutory definition of patentable subject matter: that it carves out some territory that would otherwise be covered by Section 101. Put another way, the Court can be understood to be saying that its definition *takes precedence over* the text of Section 101—that in the event of a conflict between the two, the Court’s exclusionary definition applies and Congress’s inclusionary definition does not.

This move is the first of our strategies for resolving a logical contradiction presented by competing intensional definitions of applicable doctrinal categories. Let us call it the *Trump Card*. The Trump Card move embraces the intersection between two subsets at issue, and concludes that one takes precedence over—or “trumps”—the other. For such a move to work, we must revise our model of the universe of discourse defined by the applicable legal rules. For example, we must now recognize that a process is only patentable subject matter if it is also *not* an abstract idea.

We have seen this move before. A similar logical structure undergirded our model of the law of sublease and assignment. There, we had to account for the possibility that a sublessee might assume the covenants of the primary lease, and thereby become directly liable to the primary landlord.¹²¹ Recall that in that instance, an assumption of covenants by a sublessee provided an *exception* to the rule that sublessees cannot be held directly liable. The fact that we encountered the same *formal structure* of doctrine in a body of law that we modeled via the set-theoretic operation of subtraction suggests that subtraction could be a useful tool not only for modeling settled doctrine, but for resolving contradictions among competing relevant legal authorities if and when they are first presented.

But set subtraction is not our only option for modeling the Trump Card. Formally we could express the move in another way. Let us return to the patentable subject matter example. One possible way of applying the Trump Card to this example would be to refine our *predicates* for patentable (and unpatentable) subject matter using the logical operation of conjunction with a negated premise. Thus:

¹²⁰ Mayo Collaborative Servs. v. Prometheus Labs., Inc., 566 U.S. 66, 70 (2012).

¹²¹ *Supra* Section II.B.1.

$$A = \{x: P(x) \wedge \neg W(x)\}$$

$$H = \{x: W(x)\}$$

This approach does capture the logic of the Trump Card, in that a subject matter that satisfies predicate *P* (i.e., is a process) will still be categorized as unpatentable if it also satisfies predicate *W* (i.e., is an abstract idea). However, revising our predicates in this way may be an insufficient basis for a comprehensive model of this particular body of doctrine if our intensional definitions—inclusionary and exclusionary—are not exhaustive of the universe of discourse. That is, we might still wonder whether a subject matter that is *neither* within the statute's inclusionary definitions *nor* within the judicially developed exclusionary definitions is patentable or not.

The doctrinal answer appears to be that it is not—that even if a patent claim does not fall within one of the judicial exceptions, so long as it is not within one of the statute's inclusionary categories it will still be unpatentable. Examples of such subject matter include a company, an arrangement of printed matter, or a collection of data.¹²² Accounting for this possibility in our model of the doctrine, we must imagine the relevant universe of discourse as being divided into patentable subject matter—which includes things that are within the inclusionary statutory definitions but *not* within the judicial exclusionary definitions—and unpatentable subject matter—which includes *everything else*, including *but not limited to* things that are within the judicial exclusionary definitions.¹²³ In other words, it requires us to frame our two binary outcomes—patentable versus unpatentable subject matter—in terms of the set theoretic concept of *complementation*. Patentable subject matter is defined intensionally (with a combination of inclusionary and exclusionary definitions); unpatentable subject matter is simply

¹²² For a list of “[n]on-limiting examples” of such subject matter, see 2016 *Patent Subject Matter Eligibility* [R-08.2017], U.S. PAT. & TRADEMARK OFF., <https://www.uspto.gov/web/offices/pac/mpep/s2106.html> [<https://perma.cc/92WZ-MQ7G>] (last visited Mar. 26, 2019) (§ 2106.04(c)) (collecting cases).

¹²³ Despite this *substantive* default in favor of unpatentability, as a *procedural* matter Federal Circuit doctrine stages patent examination in such a way that “the examiner bears the initial burden . . . of presenting a *prima facie* case of unpatentability.” *In re Oetiker*, 977 F.2d 1443, 1445 (Fed. Cir. 1992) (emphasis omitted). At least one commentator blames this procedural device for issuance of patents that do not meet the standards of patentability. See Sean B. Seymore, *Patent Asymmetries*, 49 U.C. DAVIS L. REV. 963, 976–91 (2016).

everything in the universe of discourse that is not patentable subject matter.¹²⁴

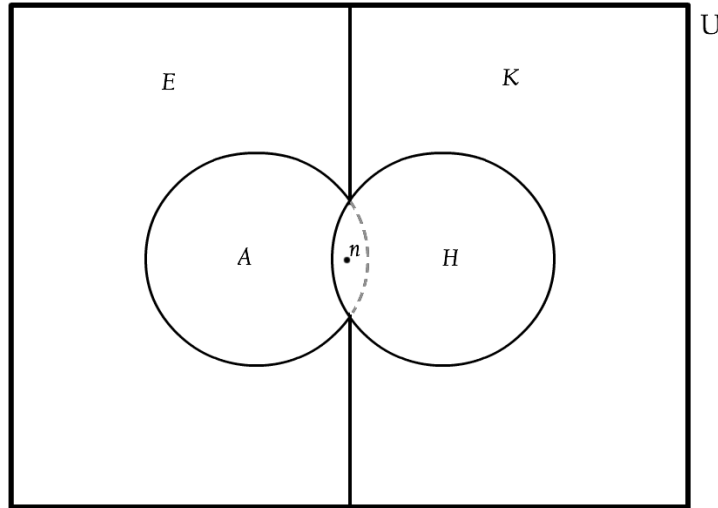
Because we must resort to complementation if we are to exhaust the universe of discourse, revising the intensional *predicates* of the sets corresponding to our legal categories is insufficient to comprehensively resolve the potential for contradiction identified earlier. Instead, we must revise the *inclusion relations* between the set of patentable subject matter and its subsets. As with the law of sublease and assignment—and as foreshadowed above—we achieve this through the operation of *subtraction*. Formally, with respect to the relationship between the set of “processes” and the set of “abstract ideas,” we achieve this as follows:

$$\begin{aligned} H &\subseteq K \\ (A - H) &\subseteq E \end{aligned}$$

Thus, the graphic model of these sets can now be represented as follows, eliminating the earlier contradiction:

¹²⁴ Even if the default position were reversed—that is, even if the Supreme Court’s exclusionary rule exhausted the category of unpatentable subject matter, leaving everything else patentable—we would still need to rely on complementation and inclusion rather than predication to arrive at an exhaustive model of the universe of discourse. Of course, if an exhaustive model is not deemed necessary or desirable—if, for example, one is willing to tolerate “gaps” in the law—predication might well be a sufficient basis for modeling the interaction of mutually inconsistent rules such as those discussed in this Section. Indeed, it is the legal positivist’s insistence on some authoritative source for legal rules in ascertainable social facts that has the potential to generate such “gaps” in doctrinal models. *See generally* Joseph Raz, *Legal Reasons, Sources, and Gaps*, in *THE AUTHORITY OF LAW: ESSAYS ON LAW AND MORALITY* 53 (2009).

Figure 16



When we expand our focus to consider all the predicates underlying our model of the doctrine of patentable subject matter, we can use subtraction to revise our inclusion relations in a similar way to resolve all the contradictions among our rules via the Trump Card move. Expressing this series of moves formally, we have:

$$\begin{aligned} (F \cup G \cup H) &\subseteq K \\ E &= (A \cup B \cup C \cup D) - (F \cup G \cup H) \\ K &= E' \end{aligned}$$

This formalization of the Trump Card explains an important and useful strategy for resolving conflicts between contradictory legal rules. But as the discussion of this Part demonstrates, arriving at this formalization required us to think hard about not only the rules themselves, nor even their application to a particular set of facts, but how they relate to one another, to various individual cases, and to the full universe of discourse. Moreover, it required us to resolve an ambiguity in our logic: We had to determine whether the relationships among the different features of our model was best captured by a modification of the predicates of certain sets (that is, the membership relation between rules and cases), or by a modification of the relationships between those sets (that is, the inclusion relation between

rules and adjudicative outcomes). We will more thoroughly examine this ambiguity in Part IV. First, let us review some alternative strategies for avoiding the logical contradictions threatened by conflicting rules.

C. *The No True Scotsman*

The No True Scotsman is an analytical move identified by philosopher Antony Flew, who illustrated it by means of an anecdote:

Imagine some Scottish chauvinist settled down one Sunday morning with his customary copy of *The News of the World*. He reads the story under the headline, 'Sidcup Sex Maniac Strikes Again'. Our reader is, as he confidently expected, agreeably shocked: 'No Scot would do such a thing!' Yet the very next Sunday he finds in that same favourite source a report of the even more scandalous on-goings of Mr Angus MacSporran in Aberdeen. This clearly constitutes a counter example, which definitively falsifies the universal proposition originally put forward. . . . Allowing that this is indeed such a counter example, he ought to withdraw; retreating perhaps to a rather weaker claim about most or some. But even an imaginary Scot is, like the rest of us, human; and we none of us always do what we ought to do. So in fact what he says is: 'No *true* Scotsman would do such a thing!'¹²⁵

The No True Scotsman move allows us to maintain our predicates and the sets constructed by them unchanged, at the cost of implausibly denying their application to certain objects we rely on them to classify. As Flew puts it, "[a] bold, indeed reckless, claim about all those who happen to be members of a certain category is being surreptitiously replaced by an utterance which is, in effect, made true by an arbitrary redefinition."¹²⁶

Many determinations in legal analysis have this kind of arbitrary feel. Wherever a doctrine is subject to a flexible standard, a multi-factor balancing test, or a totality-of-the-circumstances inquiry—as opposed to a bright-line rule—reasonable minds will frequently differ as to whether

¹²⁵ ANTONY FLEW, THINKING ABOUT THINKING 47 (1975) (emphasis added). Sidcup is a suburban neighborhood of southeastern London.

¹²⁶ *Id.* at 47–48.

marginal cases fall within or without the category defined by the applicable legal doctrine. Intellectual property law, for example, has no shortage of such fuzzy doctrines: the “likelihood of confusion” multi-factor tests for trademark infringement;¹²⁷ the “substantial similarity” standard for copyright infringement;¹²⁸ the “fair, reasonable, and nondiscriminatory” standard for patent licensing terms;¹²⁹ the “non-obviousness” standard for patent validity.¹³⁰ Such standards are perhaps even more common in less specialized areas of law: disagreements abound as to what conduct is consistent with the duty of “reasonable care,”¹³¹ what constitutes a “compelling state interest,”¹³² or what is consistent with “traditional notions of fair play and substantial justice.”¹³³

We will return to the theoretical issues underlying the rules-versus-standards dichotomy and its relation to the set-theoretic model we are developing later.¹³⁴ For now, we may confine ourselves to the observation that such standards are especially vexing when they threaten the type of doctrinal contradiction we have been examining in this Part. To illustrate the problem, let us consider the interaction of the

¹²⁷ 4 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 24:28–43 (4th ed. 1996); *Polaroid Corp. v. Polarad Elecs. Corp.*, 287 F.2d 492, 495 (2d Cir. 1961); *AMF Inc. v. Sleekcraft Boats*, 599 F.2d 341, 348–49 (9th Cir. 1979).

¹²⁸ 4 NIMMER ON COPYRIGHT § 13.03; *Arnstein v. Porter*, 154 F.2d 464, 468 (2d Cir. 1946); *Sid & Marty Krofft TV Prods. v. McDonald’s Corp.*, 562 F.2d 1157, 1164 (9th Cir. 1977).

¹²⁹ *Microsoft Corp. v. Motorola, Inc.*, 795 F.3d 1024, 1048 (9th Cir. 2015). In litigation contexts, this standard intersects with yet another complex and fuzzy standard: the 15-factor analysis for calculating patent damages under *Georgia-Pacific Corp. v. U.S. Plywood Corp.*, 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970).

¹³⁰ 35 U.S.C. § 103 (1976); *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

¹³¹ DAN B. DOBBS, PAUL T. HAYDEN & ELLEN M. BUBLICK, *THE LAW OF TORTS* § 127 (2d ed. 2011) (“The duty owed by all people generally—the standard of care—is the duty to exercise the care that would be exercised by a reasonable and prudent person under the same or similar circumstances to avoid or minimize risks of harm to others.”).

¹³² See generally Stephen A. Siegel, *The Origin of the Compelling State Interest Test and Strict Scrutiny*, 48 AM. J. LEGAL HIST. 355 (2006).

¹³³ *Int’l Shoe Co. v. Washington*, 326 U.S. 310, 316 (1945); see also *Kulko v. Superior Court*, 436 U.S. 84, 92 (1978) (“Like any standard that requires a determination of ‘reasonableness,’ the ‘minimum contacts’ test of *International Shoe* is not susceptible of mechanical application . . . We recognize that this determination is one in which few answers will be written in black and white. The greys are dominant and even among them the shades are innumerable.”) (internal quotations and citations omitted).

¹³⁴ See *infra* Section IV.B.

copyright doctrines of fair use and the derivative works right. Section 106 of the Copyright Act gives copyright owners the exclusive right “to prepare derivative works based upon the copyrighted work,” and to authorize others to do so—unauthorized preparation of a derivative work is an infringement of copyright.¹³⁵ However, this right is “subject to” the doctrine of copyright fair use—codified in Section 107 of the Act—which provides that a “fair use” of a copyrighted work “is not an infringement of copyright.”¹³⁶ And “fair use” is a notoriously imprecise doctrinal category. A finding of fair use depends on a weighing of four statutory factors, each of which derives from—and has been further developed by—a long line of case law.¹³⁷ Moreover, there appears to be significant overlap between the categories “derivative work” and “fair use.”

Section 101 of the Copyright Act defines a “derivative work” as “a work based upon one or more pre-existing works, . . . [in any] form in which a work may be recast, *transformed*, or adapted.”¹³⁸ But under current interpretations of Section 107, perhaps the single most important factor in determining whether a use is “fair” (and therefore non-infringing) is whether it is “*transformative*”: that is, whether it “adds something new, with a further purpose or different character, altering the first [work] with new expression, meaning, or message.”¹³⁹ Indeed, Jane Ginsburg recently concluded (or perhaps lamented) that under current doctrine, “if the [defendant’s] use is ‘transformative,’ the four-factor statutory test [for fair use under Section 107] effectively reduces to a single factor.”¹⁴⁰ Thus, the “transformation” of a

¹³⁵ 17 U.S.C. §§ 106(2) (2002), 501(a) (2002).

¹³⁶ 17 U.S.C. §§ 106–07 (1976).

¹³⁷ 17 U.S.C. § 107 (1976); 4 WILLIAM F. PATRY, PATRY ON COPYRIGHT § 10 (2007).

¹³⁸ 17 U.S.C. § 101 (2010) (emphasis added).

¹³⁹ *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578–79 (1994) (citing Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1111 (1990)).

¹⁴⁰ Jane C. Ginsburg, *Fair Use for Free, or Permitted-but-Paid?*, 29 BERKELEY TECH. L.J. 1383, 1400 (2014); *but see* Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L. REV. 549, 595, 604–05 (2008) (“It appears . . . that courts and commentators have exaggerated the influence of transformativeness doctrine on our fair use case law. . . . Nevertheless, in those opinions in which transformativeness did play a role, it exerted nearly dispositive force not simply on the outcome of factor one but on the overall outcome of the fair use test. More specifically, the data suggest that while a finding of transformativeness is not necessary to trigger an overall finding of fair use, it is sufficient to do

copyrighted work without authorization may be both a necessary and a sufficient basis to categorize a particular use of a copyrighted work as *both* an infringing derivative work under Section 106 *and* a non-infringing fair use under Section 107. This contradiction has attracted the attention of copyright scholars.¹⁴¹

Note that this contradiction is *structurally* more difficult to resolve via the Trump Card than was the patentable subject matter contradiction. In the context of patentable subject matter, we had to determine how to categorize an invention that was—for example—both a “process” and an “abstract idea.” Here, we have to determine how to categorize a use that is both a “transformation” for purposes of the derivative works right and “transformative” for purposes of fair use doctrine. In the latter context, it is not clear from the *natural language* of the relevant predicates that the two doctrinal subsets of interest are in fact different sets. If they are not, then we would be unable to avoid the self-contradictory conclusion that a use of a copyrighted work that falls into this set is simultaneously a derivative-works-right infringement and a non-infringing fair use simply by subtracting one doctrinal set from the other. (Subtraction would yield only the null set, implying that the one doctrine completely abrogates the other.)

So let us assume for the moment that “uses that transform a copyrighted work” and “uses that are transformative of a copyrighted work” are in fact nonequivalent sets, and that all transformative uses are fair while all uses that transform are infringing. We can begin to model these sets along the lines of the model developed earlier in the context of patentable-subject-matter doctrine (though for now the model will be much simpler):

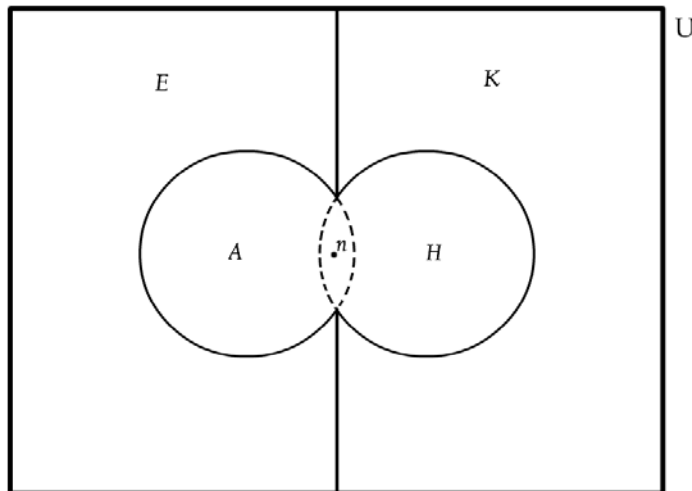
so.”). Professor Beebe’s regression analysis revealed that no factual finding had a stronger association with a finding of fair use than a finding of transformativeness, and that only one factual finding—that the defendant took the “heart” of the plaintiff’s copyrighted work—had a stronger association in favor of either party.

¹⁴¹ See generally, e.g., R. Anthony Reese, *Transformativeness and the Derivative Work Right*, 31 COLUM. J.L. & ARTS 467 (2008).

$$\begin{aligned}
 E &= \{x: x \text{ is an infringement of copyright}\} \\
 K &= \{x: x \text{ is not an infringement of copyright}\} \\
 P(x) &= x \text{ is a use that transforms a copyrighted work} \\
 A &= \{x: P(x)\} \\
 A &\subseteq E \\
 W(x) &= x \text{ is a transformative use of a copyrighted work} \\
 H &= \{x: W(x)\} \\
 H &\subseteq K
 \end{aligned}$$

Again, it seems quite likely that sets A and H would have a large area of intersection, but to serve our purposes in modeling the strategies for resolving logical contradictions in doctrine we can represent our model of copyright doctrines graphically as *formally* identical to our model of patentable subject matter doctrine, to highlight the structural similarities (and differences) between the two strategies we are examining. We complete the model by positing some case n that is both “a transformative use” and “a use that transforms,” and represent our universe of discourse graphically (yielding a familiar picture):

Figure 17



The No True Scotsman is a strategy to resolve this contradiction by positing that n is not in fact a “transformative use,” even though it would seem to be under conventional understandings of the natural

language of the predicate. The copyright example illustrates the strategy. In much fair use litigation, the key question is whether the defendant's use is sufficiently "transformative" to qualify as a fair use, even though it is undisputed that the defendant's use has in some way changed the plaintiff's work in making a new use of it. Diane Zimmerman described the underlying dilemma, and the moves by which courts resolve it:

What *is* hard to understand is why the courts engaged in so much twisting and turning to avoid the seemingly obvious conclusion that, whatever else might have been troubling in the defendants' cases, the uses in question were at least "transformative": they clearly did provide the public with a new or substantially reworked product. . . . Judges who face a kind of transformative use that they strongly believe ought to be controlled by the plaintiff alone may well think that they are caught between the proverbial rock and a hard place. It is little wonder, therefore, that they sometimes resolve their dilemma by performing a little deft Lewis Carroll-type surgery on the inconvenient word to create an escape for themselves from that uncomfortable space.¹⁴²

We can identify this "twisting and turning" or "Lewis Carroll-type surgery" as an instance of the No True Scotsman in action. For example, in *Castle Rock Entertainment, Inc. v. Carol Publishing Group, Inc.*¹⁴³—cited by Professor Zimmerman¹⁴⁴—the Second Circuit affirmed then-District Judge Sotomayor's award of summary judgment to the owners of the copyright in the popular television show *Seinfeld* against the publisher of an unauthorized book of trivia based on the show. The court cited Section 101's definition of "derivative work," but explained: "Although derivative works that are subject to the author's copyright transform an original work into a new mode of presentation, such works—unlike works of fair use—take expression for purposes that are

¹⁴² Diane Leenheer Zimmerman, *The More Things Change the Less They Seem "Transformed": Some Reflections on Fair Use*, 46 J. COPYRIGHT SOC'Y. U.S. 251, 259–60 (1999) (emphasis in the original). The reference to Lewis Carroll invokes his character of Humpty-Dumpty, who insisted: "When *I* use a word, it means just what I choose it to mean—neither more nor less." *Id.* at 252 n.5 (quoting LEWIS CARROLL, *THROUGH THE LOOKING GLASS—AND WHAT ALICE FOUND THERE* 94 (1946)).

¹⁴³ 150 F.3d 132 (2d Cir. 1998).

¹⁴⁴ Zimmerman, *supra* note 142, at 251–52 & n.2.

not ‘transformative.’”¹⁴⁵ More recently, in *Cariou v. Prince*, the Second Circuit considered thirty works in which appropriation artist Richard Prince reproduced and then modified copyrighted works by photographer Patrick Cariou without permission. The court found twenty-five of these thirty works to be fair uses as a matter of law, on grounds that “Prince’s composition, presentation, scale, color palette, and media are fundamentally different and new compared to the photographs, as is the expressive nature of Prince’s work.”¹⁴⁶ But—citing *Castle Rock* on the relationship between derivative works and fair use¹⁴⁷—it remanded for further findings regarding the other five works.¹⁴⁸ Its reasoning is an encapsulation of the No True Scotsman at work: “[W]e cannot say for sure . . . whether Prince has transformed Cariou’s work enough to render it transformative.”¹⁴⁹

The twisting and turning language in these cases does not seem to be invoking the Trump Card—that is, these courts do not hold that if a defendant’s use is transformative, it is *perforce* not a derivative work, or that a defendant’s use that transforms the plaintiff’s work is *perforce* not fair.¹⁵⁰ Instead, many courts appear to treat the category of derivative works and the category of fair uses as if they have no determinate relationship to one another—and in doing so are able to superficially avoid the contradiction presented in our formal model. The No True Scotsman allows them to do so by working at the level of the individual

¹⁴⁵ *Castle Rock*, 150 F.3d at 143.

¹⁴⁶ *Cariou v. Prince*, 714 F.3d 694, 706–10 (2d Cir. 2013).

¹⁴⁷ *Id.* at 708.

¹⁴⁸ *Id.* at 710–11.

¹⁴⁹ *Id.* at 711.

¹⁵⁰ Indeed, one district court that held as much was criticized by Professor Reese as “clearly incorrect.” Reese, *supra* note 141, at 469–70, quoting *Clean Flicks of Colo., LLC v. Soderbergh*, 433 F. Supp. 2d 1236, 1242 (D. Colo. 2006) (“[B]ecause the infringing copies of these movies are not used in a transformative manner, they are not derivative works and do not violate § 106(2).”) (internal quotations omitted). The Seventh Circuit has similarly criticized Second Circuit caselaw on grounds that it commits the same error. *Kienitz v. Sconnie Nation LLC*, 766 F.3d 756, 758 (7th Cir. 2014), *cert. denied*, 135 S. Ct. 1555 (2015) (“To say that a new use transforms the work is precisely to say that it is derivative and thus, one might suppose, protected under § 106(2). *Cariou* and its predecessors in the Second Circuit do not explain how every ‘transformative use’ can be ‘fair use’ without extinguishing the author’s rights under § 106(2).”). As we will see, however, the Seventh Circuit is confusing the *substance* of the Second Circuit’s approach to this particular doctrinal problem with the *structure* of that approach. See *infra* Part IV.

uses at issue in each case, finding that the defendant's conduct falls within one of the two conflicting categories and summarily asserting it does not fall within the other. For example, the court in *Micro Star v. FormGen Inc.* provided a lengthy analysis to support its conclusion that the defendant had created an unauthorized derivative work by creating additional levels for plaintiff's copyrighted video game, and then proceeded to dismiss the question of transformativeness in a single sentence buried in a footnote.¹⁵¹ Similarly, the court in *Nihon Deizai Himbun, Inc. v. Comline Business Data, Inc.*—in which the defendant had prepared English-language abstracts of plaintiff's Japanese-language articles—the court performed a thorough analysis of plaintiff's derivative-works-right claim, then in two sentences concluded that the defendant's uses were “not in the least transformative” because they were “for the most part direct translations . . . [that] added almost nothing new.”¹⁵² In both cases, the courts' assertion that the defendants' uses of the plaintiffs' works were not “transformative” seems inconsistent with the ordinary meaning of the term. But as *Castle Rock* suggests with its use of scare quotes, just because a derivative work “transforms” the plaintiff's expression, that does not make such a work “transformative” as a legal matter.¹⁵³

Such a distinction between the ordinary meaning and the legal meaning of a particular term as applied to a particular case is a hallmark of the No True Scotsman. Rhetorically at least, such distinctions are fairly common—indeed, they are often signaled with scare quotes—and they help to resolve a case without disturbing a model of the universe of discourse that depends on the disjointness of two sets at issue in the case. The earlier history of the doctrine we investigated in the previous section—the law of patentable subject matter—provides additional examples of this dynamic. In *Parker v. Flook*, for instance, the Court conceded that the applicant's method “is a ‘process’ in the ordinary sense of the word,”¹⁵⁴ but also noted that the Court's earlier holding in

¹⁵¹ The entirety of the court's analysis of the transformativeness of the defendant's use was: “[It] can hardly be described as transformative; anything but.” *Micro Star v. FormGen Inc.*, 154 F.3d 1107, 1113 n.6 (9th Cir. 1998).

¹⁵² *Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc.*, 166 F.3d 65, 72 (2d Cir. 1999) (internal quotations omitted).

¹⁵³ See *supra* text accompanying notes 143–45.

¹⁵⁴ *Parker v. Flook*, 437 U.S. 584, 588 (1978).

*Gottschalk v. Benson*¹⁵⁵ “forecloses a purely literal reading of § 101.”¹⁵⁶ Indeed, in *Benson* the Court had telegraphed the availability of the No True Scotsman move, framing the issue in that case as “whether the method described and claimed is a ‘process’ *within the meaning of the Patent Act.*”¹⁵⁷ (The fact that these opinions used language suggesting a No True Scotsman in a universe of discourse that we ultimately identified with the Trump Card presents interesting theoretical issues that will be explored below.)

In both the older patentable subject matter cases and the more recent fair use cases, courts are distinguishing between the *ordinary* meaning of certain intensional definitions and the *legal* meaning of those definitions *in application to a particular state of the world*. In other words, they are asking whether *this* process is a *true* “process,” or *this* transformation is *truly* “transformative.” In formal terms, the No True Scotsman can be represented in our model of the universe of discourse with a single move—evaluating the troublesome predicate in such a way as to remove the contradictory membership relation:

$$\neg W(n) \\ \therefore n \notin H$$

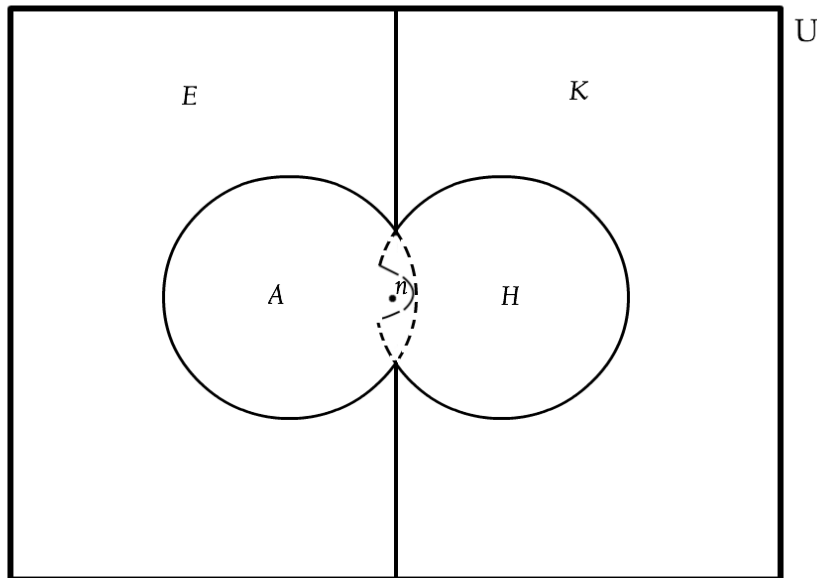
Graphically, the move can be depicted thus:

¹⁵⁵ 409 U.S. 63 (1972).

¹⁵⁶ *Flook*, 437 U.S. at 588–89.

¹⁵⁷ *Benson*, 409 U.S. at 64 (emphasis added).

Figure 18



The No True Scotsman is a strategy that releases the immediate pressure of an apparent contradiction, but in doing so only increases the strain on the logical structure that created the contradiction in the first place. As cases that seem on their face to fall within the natural-language intensional predicate of one doctrinal set are placed outside that set by particular acts of *ad hoc* adjudication, the coherence of doctrinal sets across time and circumstance becomes difficult to maintain—the link between set *predication* and set *membership* begins to break down. Because law is a human institution, such incoherence can endure for a surprisingly long time without stimulating any radical reorganization of set-theoretic models. Indeed, Flew predicted as much in identifying the No True Scotsman as a feature of human psychology. Recall his original example: The case of Agnus MacSporran should probably have led the Scottish chauvinist to reevaluate the role of the category “Scotsmen” in his model of the relevant universe of discourse. In particular, it ought to have led him to reconsider whether members of this category all share the distinctive characteristics he ascribes to them (i.e., whether the category “Scotsmen” has a nonempty intersection with the category of

“people who would do such a thing”). But Flew thinks such reevaluation is unlikely—that we are resistant to revising our set-theoretic models of the world around us if we can avoid it. As he puts it: “The temptation of course is not just to slide, under the pressure of falsification, . . . [i]t is to fail to recognize what has happened, and so to be apt to slide back again into the original interpretation immediately that pressure is removed.”¹⁵⁸

This psychological conjecture has some empirical support,¹⁵⁹ and it is a particularly salient concern for legal analysis, reflected in the old adage that hard cases make bad law. It is precisely when the categorization of a set of facts as outside some seemingly applicable rule (e.g., Agnus MacSporran is no *true* Scotsman) is inconsistent with conventional understandings of the natural language in which that rule is framed (e.g., Agnus MacSporran is a native Aberdonian; Aberdeen is in Scotland) that we run the risk of inconsistency in our set-theoretic models. This inconsistency, as Flew’s example suggests, may often arise from results-oriented decision-making, and it can thus pose a direct threat to the rule of law.¹⁶⁰ But the common law has a built-in check on this kind of free-wheeling ad hoc decision-making: the recursive process of extensional set-building that we earlier identified with the institution

¹⁵⁸ FLEW, *supra* note 125, at 50. Similar tendencies have been flagged in philosophical debates under the name of “humpty-dumptying” or “motte-and-bailey doctrines.” See generally Nicholas Shackel, *The Vacuity of Postmodernist Methodology*, 36 METAPHILOSOPHY 295 (2005).

¹⁵⁹ See generally, e.g., Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULL. 480 (1990); Dan M. Kahan, *Ideology, Motivated Reasoning, and Cognitive Reflection: An Experimental Study*, 8 JUDGMENT & DECISION MAKING 407 (2012); Ziva Kunda & Lisa Sinclair, *Motivated Reasoning with Stereotypes: Activation, Application, and Inhibition*, 10 PSYCHOL. INQUIRY 12 (1999).

¹⁶⁰ The rule of law is obviously a complex and contested concept, but for present purposes it simply refers to some standard of rationality and consistency in the application of legal rules. For more on the concept, see generally Richard H. Fallon, Jr., “*The Rule of Law*” as a Concept in Constitutional Discourse, 97 COLUM. L. REV. 1 (1997); cf. Dworkin, *supra* note 76, at 1064 (“Judges, like all political officials, are subject to the doctrine of political responsibility. This doctrine states, in its most general form, that political officials must make only such political decisions as they can justify within a political theory that also justifies the other decisions they propose to make. The doctrine seems innocuous in this general form; but it does, even in this form, condemn a style of political administration that might be called, following Rawls, intuitionistic. It condemns the practice of making decisions that seem right in isolation, but cannot be brought within some comprehensive theory of general principles and policies that is consistent with other decisions also thought right.”).

of precedent.¹⁶¹ This process is key to the final contradiction-avoidance strategy we must review.

D. *The Tertium Quid*

The No True Scotsman ties the disposition of a case to the specific facts of that case. In so doing, it purports to limit the effect of the adjudication of the case on the model of the applicable doctrinal universe of discourse—and particularly on the predicate for the set from which the case is to be excluded. But because the *outcome* of a common-law adjudication has precedential value independent of the announced rationale for that outcome, such efforts to limit a case to its facts are inherently unstable. As Professor Schauer notes, it may be that law, as an institution, inherently requires outcomes to be justified with reasons that extend beyond the facts of a single case.¹⁶² Thus, overt efforts by courts to limit cases to their facts generally fail.

In intellectual property law we see an example of this dynamic in the history of the doctrine of common-law misappropriation. A federal common-law claim for misappropriation was recognized by the Supreme Court in *International News Service v. Associated Press*,¹⁶³ in which a wire service sued its competitor for copying its news stories from East Coast newspapers and selling them to West Coast newspapers at a discount. The Court held that the defendant could be held liable for “appropriating to itself the harvest of those who have sown,”¹⁶⁴ dismissing any potential contradiction with the Copyright Act’s exclusion of factual material from exclusive ownership with a wave of its hand: “We need spend no time . . . upon the general question of property in news matter at common law, or the application of the copyright act, since it seems to us the case must turn upon the question of unfair competition in business.”¹⁶⁵

Recognizing that this doctrinal contradiction could not be so casually dismissed, Judge Learned Hand undertook in the case of

¹⁶¹ See *supra* Section II.C.

¹⁶² Frederick Schauer, *Giving Reasons*, 47 STAN. L. REV. 633, 653 & n.60 (1995).

¹⁶³ 248 U.S. 215 (1918).

¹⁶⁴ *Id.* at 239–40.

¹⁶⁵ *Id.* at 234–35.

Cheney Bros. v. Doris Silk Corp. to limit *INS v. AP* to its facts.¹⁶⁶ Judge Hand treated *INS v. AP* as an example of a No True Scotsman—“where the occasion is at once the justification for, and the limit of, what is decided.”¹⁶⁷ But his approbation notwithstanding, the misappropriation cause of action continues to be asserted in the Second Circuit to this day. In *National Basketball Ass’n v. Motorola, Inc.*,¹⁶⁸ the court listed the elements of the cause of action in general terms—deriving them primarily from the *facts* of *INS*.¹⁶⁹ And those elements have now become the basis for a family of cases in which claims are asserted for misappropriation of “hot news”—the category of information that was found to have been misappropriated in *INS*.¹⁷⁰

This tendency of individual cases to become the seeds of intensionally-defined doctrinal sets—even where courts go to great lengths to prevent them from doing so—is an inevitable consequence of the recursive nature of the common law. As discussed in Section II.C above, doctrinally relevant sets may be constructed not only intensionally, based on predicates found in the language of statutes or judicial opinions, but extensionally, by collating and reconciling the facts and results of individual cases that have precedential value. And in areas of doctrine that tend to generate No True Scotsmen—areas where existing intensional definitions conflict to the point where courts are categorizing cases within or without those definitions on a seemingly *ad hoc* basis—the recursive process of cataloguing and comparing the facts and outcomes of these cases tends toward the development of some *new* intensional definition emergent from the cases themselves. We saw an example of this process in *Kremen*’s effort to reconcile two inconsistent definitions of “property” for purposes of California conversion law.¹⁷¹ Such a newly developed intensional principle is necessarily distinct from any pre-existing intensional definitions, which are therefore admitted to

¹⁶⁶ *Cheney Bros. v. Doris Silk Corp.*, 35 F.2d 279, 280 (2d Cir. 1929).

¹⁶⁷ *Id.*

¹⁶⁸ 105 F.3d 841 (2d Cir. 1997).

¹⁶⁹ *See id.* at 852.

¹⁷⁰ *See, e.g., Barclays Capital Inc. v. Theflyonthewall.com, Inc.*, 650 F.3d 876 (2d Cir. 2011). *Dow Jones & Co. v. Real-Time Analysis & News, Ltd.*, No. 14 CIV. 131 JMF GWG, 2014 WL 4629967 (S.D.N.Y. Sept. 15, 2014) (default judgment), *report and recommendation adopted*, 2014 WL 5002092 (S.D.N.Y. Oct. 7, 2014).

¹⁷¹ *See supra* notes 87–96 and accompanying text.

provide an incomplete—or mistaken—model of the universe of discourse. The identification of—and reliance on—such a distinct and novel intensional principle is the last strategy for resolving doctrinal conflict. I will refer to it as the *Tertium Quid*.

The *Tertium Quid* is an old work horse of the Langdellian case method. It is one of the first tricks picked up by the first-year American law student: to distill and generalize from the facts and outcomes of an apparently inconsistent set of cases some new abstract principle—distinct from the principles explicitly relied on in the opinions themselves—to govern novel future cases.¹⁷² In the context of our model of doctrinal conflict, a hard-to-justify classification of a case as within or without an existing intensionally-defined set may suggest the need for a *new* set not accounted for in the existing model of the universe of discourse. Moreover, this new set may turn out to be *more* relevant to organizing the universe of discourse into the two universe-exhausting complementary sets we assign to binary adjudicative outcomes. Thus, in a common-law system, every No True Scotsman is the seed of a *Tertium Quid*, and every *Tertium Quid* has the capacity to supplant previously applicable doctrinal sets.

The areas of intellectual property law we have already examined each provide some examples of efforts to deploy the *Tertium Quid*. In patentable subject matter, for example, the “product of nature” exclusion has been criticized as insufficient to explain the outcome of cases such as *Myriad*.¹⁷³ As Dan Burk points out, the predicate “natural phenomenon” does not provide any principled basis for the *Myriad* Court’s distinction between a molecule of gDNA removed from its position in a human chromosome and a corresponding, exon-only cDNA molecule created in a lab. Neither molecule would have the same physical qualities as the corresponding sequence of nucleotides found *in vivo* within in a human cell.¹⁷⁴ Burk thus concludes that the “product of

¹⁷² See, e.g., Todd D. Rakoff & Martha Minow, *A Case for Another Case Method*, 60 VAND. L. REV. 597, 598–99 (2007) (“[S]tudents would be expected to work not only from the particular to the general, but also from the general to the particular.”).

¹⁷³ The *Myriad* Court itself alternately justified its holding under the “law of nature” exception, *Ass’n of Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576, 591 (2013), and the “product of nature” exception, *id.* at 595.

¹⁷⁴ Dan L. Burk, *Edifying Thoughts of a Patent Watcher: The Nature of DNA*, 60 UCLA L. REV. DISC. 92, 98–101 (2013).

nature” exclusion from patentable subject matter is generally unhelpful: “At its endpoints, the doctrine either proves everything or proves nothing. Either everything is a product of nature—drawn from and existing in the world—or nothing is a product of nature—having been intellectually and socially constructed by human cognition.”¹⁷⁵ He argues that instead, courts should—and perhaps *do*—resolve patentable subject matter cases by reference to an anti-preemption principle: that “the basic tools of scientific and technological work” are not patentable.¹⁷⁶ This anti-preemption criterion is a *Tertium Quid*: a new predicate that supposedly offers a superior basis for the mapping of cases to binary outcomes within the universe of discourse.

The indeterminacy generated by the contradiction between fair use doctrine and the derivative works right in copyright has similarly spawned efforts to deploy the *Tertium Quid*. Anthony Reese, for example, cites various Court of Appeals opinions to argue that “transformativeness” does not refer to transformation of a *copyrighted work*, but rather to the use of such a work for a *different purpose* than the purpose for which the copyright owner uses it.¹⁷⁷ Under this new intensional definition of the relevant doctrinal set, “even making an exact copy of a work may be transformative so long as the copy serves a different function than the original work.”¹⁷⁸ Of course, the same indeterminacy that makes the *Tertium Quid* attractive makes it extremely useful in contesting any new model of the universe of discourse. Professor Ginsburg, for example, has proposed dividing up the universe of transformative uses even more finely—to exclude from the fair use defense transformative purposes that are not *authorial* purposes (i.e., where the transformative purpose is to drive a new business model, such as online search).¹⁷⁹

To formalize the *Tertium Quid*, let us turn to an example from another branch of intellectual property law: trademarks. Courts have

¹⁷⁵ *Id.* at 101.

¹⁷⁶ *Id.* at 102 (quoting *Gottschalk v. Benson*, 409 U.S. 63, 67 (1972)). A similar line of analysis suggests that the Supreme Court’s “abstract ideas” jurisprudence is being subsumed by the search for an “inventive concept.” See, e.g., Mark D. Janis & Timothy R. Holbrook, *Patent-Eligible Processes: An Audience Perspective*, 17 VAND. J. ENT. & TECH. L. 349, 377–83 (2015).

¹⁷⁷ Reese, *supra* note 141, at 485.

¹⁷⁸ *Id.* (quoting *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1165 (9th Cir. 2007)).

¹⁷⁹ See generally Ginsburg, *supra* note 140.

long struggled with the application of trademark liability to defendants who are using a plaintiff's trademarks—at least in part—for expressive purposes. Rochelle Dreyfuss summarized the dilemma in a 1989 article:

[Traditionally, trademark] claims focused on the impact of the mark on purchasing decisions. By the same token, defenses centered on the commercial requirements of the competitive marketplace. The terms that delimited the reach of trademark law—consumer confusion, gap bridging, fair use, genericity, abandonment—were understood strictly by reference to these commercial interests. But as trademark owners have begun to capitalize on the salience of these symbols in the culture, the justifications that formerly delineated the scope of the law have lost significance. . . . Lacking the traditional analytical tools provided by trademark law, courts have lately attempted to apply first amendment jurisprudence to such claims. . . . [A]lthough the Constitution supplies a normative principle favoring public access to the tools of expression, the body of law that has developed under the first amendment provides a surprisingly uncongenial framework for analysis.¹⁸⁰

The basic problem of doctrinal structure is a familiar one: In some cases, the rule applying liability for trademark infringement contradicts the rule insulating expression from legal liability. Specifically, liability for infringement attaches when a defendant uses a plaintiff's trademark in a way that is "likely to cause confusion . . . as to the affiliation, connection, or association of [defendant] with [plaintiff], or as to the origin, sponsorship, or approval of [the defendant's] goods, services, or commercial activities by [plaintiff]."¹⁸¹ But the First Amendment generally forbids imposing legal liability based on the content of one's expression.¹⁸² In cases where a defendant uses a trademark in an expressive way, and thereby causes confusion, we have our familiar contradiction.

¹⁸⁰ Rochelle Cooper Dreyfuss, *Expressive Genericity: Trademarks as Language in the Pepsi Generation*, 65 NOTRE DAME L. REV. 397, 398 (1989).

¹⁸¹ 15 U.S.C. § 1125(a)(1)(A) (2018).

¹⁸² See, e.g., *R.A.V. v. City of St. Paul*, 505 U.S. 377, 382 (1992).

Once again, we can model this contradiction both formally and graphically. Formally, our universe of discourse includes the following definitions:

$$A = \{x: x \text{ is a use of a trademark that is likely to cause consumer confusion}\}$$

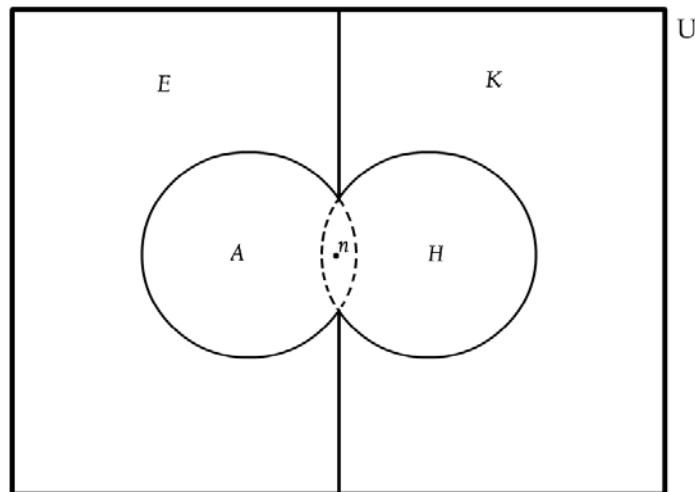
$$H = \{x: x \text{ is a form of constitutionally protected expression}\}$$

$$A \subseteq E$$

$$H \subseteq K$$

Graphically, we can again represent our model of trademark doctrines as *structurally* identical to our model of patentable subject matter doctrine and copyright fair use doctrine. We again complete the model by positing some case n in which the defendant has engaged in conduct that is both “a use of a trademark that is likely to cause consumer confusion” and “a form of constitutionally protected expression,” and arrive at the now-familiar picture of our universe of discourse:

Figure 19



While we might expect that a conflict between a constitutional provision and a statutory cause of action would ordinarily be resolved

via the Trump Card,¹⁸³ courts deciding trademark cases have not consistently invoked that strategy when faced with First Amendment defenses. Instead, for some time, they resorted to an implicit balancing analysis—weighing the seriousness of the confusion caused by the defendant’s activity against the importance of the speech interest underlying that activity. The resolution of this balancing analysis often invoked the No True Scotsman—a finding for the plaintiff or defendant based on a contestable assertion regarding the case’s satisfaction of one of the two applicable predicates. That is, in some cases the application of trademark liability was found not to impose serious burdens on expression;¹⁸⁴ in others the expression at issue was deemed unlikely to cause any kind of confusion.¹⁸⁵

¹⁸³ See U.S. CONST. art. VI, cl. 2 (“This Constitution . . . shall be the supreme Law of the Land”); *Marbury v. Madison*, 5 U.S. 137, 180 (1803) (“[A] law repugnant to the Constitution is void.”).

¹⁸⁴ *San Francisco Arts & Athletics, Inc. v. U.S. Olympic Comm.*, 483 U.S. 522, 539–40 (1987) (“The possibility for confusion as to sponsorship is obvious . . . [and t]he application of the Act to this commercial speech is not broader than necessary to protect the legitimate congressional interest and therefore does not violate the First Amendment.”); *Mut. of Omaha Ins. Co. v. Novak*, 836 F.2d 397, 402 (8th Cir. 1987) (“[T]he protection afforded by the First Amendment does not give Novak license to infringe the rights of Mutual. . . . Other avenues for Novak to express his views exist and are unrestricted by the injunction.”); *Anheuser-Busch, Inc. v. Balducci Publ’ns*, 28 F.3d 769, 776 (8th Cir. 1994) (“Balducci’s ad parody was likely to confuse consumers as to its origin, sponsorship or approval. This confusion might have to be tolerated if even plausibly necessary to achieve the desired commentary—a question we need not decide. In this case, the confusion is wholly unnecessary to Balducci’s stated purpose.”); *Dallas Cowboys Cheerleaders, Inc. v. Pussycat Cinema, Ltd.*, 604 F.2d 200, 205–06 (2d Cir. 1979) (“The public’s belief that the mark’s owner sponsored or otherwise approved the use of the trademark satisfies the confusion requirement. . . . Plaintiff’s trademark is in the nature of a property right, and as such it need not yield to the exercise of First Amendment rights under circumstances where adequate alternative avenues of communication exist.”) (internal citations and quotation marks omitted); cf. *MGM-Pathe Commc’ns Co. v. Pink Panther Patrol*, 774 F. Supp. 869, 877 (S.D.N.Y. 1991) (“[T]he harm to the plaintiff resulting from the . . . likelihood of public confusion . . . could seriously impair the value and continued usefulness of its mark. The Patrol contends MGM’s suit is barred by the First Amendment. They contend that because the Patrol is engaged in political speech, it is less subject to the trademark laws. There is no legal support for this position. The seriousness and virtue of a cause do not confer any right to the use of the trademark of another.”).

¹⁸⁵ *Lucasfilm Ltd. v. High Frontier*, 622 F. Supp. 931, 934 (D.D.C. 1985) (“Defendants use star wars in the body of their message in a descriptive manner to communicate ideas, rather than to create confusion as to sponsorship.”); *Girl Scouts of U.S. v. Personality Posters Mfg. Co.*, 304 F. Supp. 1228, 1231 (S.D.N.Y. 1969) (“Even if we hypothesize that some viewers might at first blush believe that the subject of the poster is actually a pregnant Girl Scout, it is highly

Eventually, however, this type of *ad hoc* adjudication—and its reliance on the No True Scotsman—came to be replaced by the Tertium Quid. In *Rogers v. Grimaldi*, the Second Circuit imposed a new structure on the analysis applicable to the unauthorized use of trademarks in certain expressive contexts:

We believe that in general the [Lanham] Act should be construed to apply to artistic works only where the public interest in avoiding consumer confusion outweighs the public interest in free expression. In the context of allegedly misleading titles using a [protectable trademark], that balance will normally not support application of the Act unless the title has no artistic relevance to the underlying work whatsoever, or, if it has some artistic relevance, unless the title explicitly misleads as to the source or the content of the work.¹⁸⁶

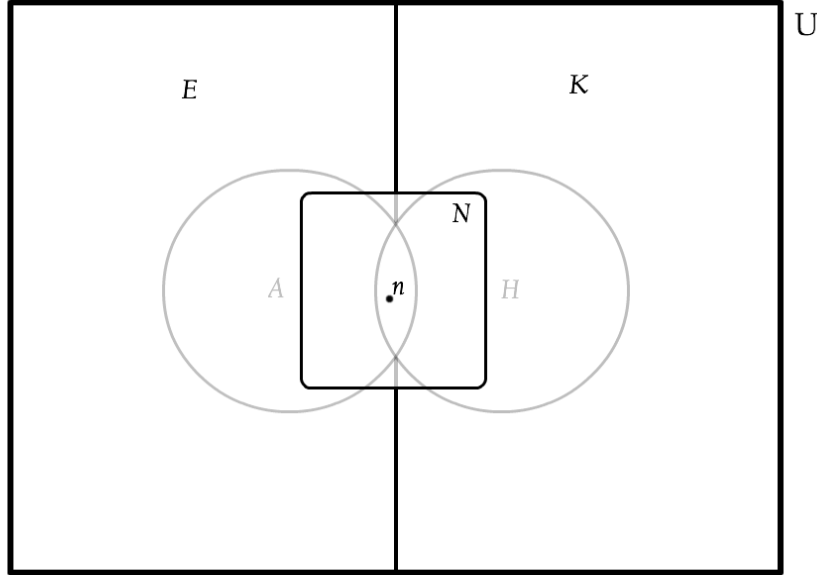
The *Rogers* test explicitly acknowledged a category of cases in which the two contradictory doctrinal sets may intersect—cases involving artistic works whose titles include someone else’s trademark—and constructed a new model of that doctrinal space. The most important implication of this move is to suggest that the pre-existing model of the universe of discourse—and in particular the intensionally defined sets that had been thought to map to binary outcomes—are no longer helpful in determining at least some relevant membership relations between cases and outcomes. The identification of a new intensional definition of a relevant set of cases, which intersects at least in part with the

doubtful that any such impression would be more than momentary or that any viewer would conclude that the Girl Scouts had printed or distributed the poster.”); *cf.* *L.L. Bean, Inc. v. Drake Publishers, Inc.*, 811 F.2d 26, 33 (1st Cir. 1987) (“The district court’s application of the Maine anti-dilution statute to appellant’s noncommercial parody cannot withstand constitutional scrutiny. Drake has not used Bean’s mark to identify or market goods or services; it has used the mark solely to identify Bean as the object of its parody.”).

¹⁸⁶ *Rogers v. Grimaldi*, 875 F.2d 994, 999 (2d Cir. 1989). The mark in question in *Rogers* was the name of the famous dancer Ginger Rogers, which she claimed was being unfairly used in the title of the Federico Fellini film “Ginger and Fred.” Rogers asserted that the title gave rise to a false endorsement claim under the Lanham Act, in essence claiming her name served as a protectable trademark—a claim that has been recognized by the federal courts in other contexts. See 5 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 28:15 (5th ed. 2018) (“Courts hold that in the context of § 43(a)(1)(A), 15 U.S.C.A. § 1125(a)(1)(A), a human persona or identity is a kind of ‘trademark’ which is infringed by a false endorsement.”) (citing cases).

intersection of the pre-existing contradictory sets, is the defining feature of the Tertium Quid. Graphically, we can model the move as follows:

Figure 20



The mere assertion of the existence of this new relevant set N does not in itself provide any information about the relationship between N and the binary outcomes in the universe of discourse. That is, the Tertium Quid does not imply any *particular* relationship between the new set N and the ultimate complementary sets E and K —only that there exists *some* such relationship (for at least *some subset* of N), and that this relationship takes precedence over the inclusion relationships between A and E on the one hand and H and K on the other.¹⁸⁷

¹⁸⁷ Formally, we can represent the logic of this move as follows:

$$\exists(N)[N \cap (A \cap H) \neq \emptyset]$$

$$n \in N$$

$$\exists(Z)[(Z \subseteq N) \wedge ([Z \subseteq E] \oplus [Z \subseteq K])]$$

For an example of the ways such a newly identified set can be mapped to binary outcomes, we can return to *Rogers*. Within the new set *Rogers* had identified (i.e., titles of expressive works that contain someone else's trademark), the court defined two new subsets and established their relations to the ultimate complementary sets defining the binary outcomes of interest. Each of these two subsets—cases in which the title has no artistic relevance to the underlying work, and cases in which the title explicitly misleads as to the source or content of the work—were then included within the set corresponding to liability; by implication expressive works that were not members of either subset were immune from liability. “Ginger and Fred” was found not to fall within either of the identified subsets, and therefore not to subject its producers to liability.¹⁸⁸ Formally we may model the analysis as follows:

$$N = \{x: x \text{ is an artistic work whose title contains someone else's trademark}\}$$

$$L = \{x: x \text{ is an artistic work in which the inclusion of someone else's trademark in the title has no artistic relevance to the underlying work}\}$$

$$M = \{x: x \text{ is an artistic work in which the inclusion of someone else's trademark in the title explicitly misleads as to the source or content of the work}\}$$

$$L \subseteq N$$

$$M \subseteq N$$

$$(L \cup M) \subseteq E$$

$$N - (L \cup M) \subseteq K$$

$$n \in N$$

$$n \notin L$$

$$n \notin M$$

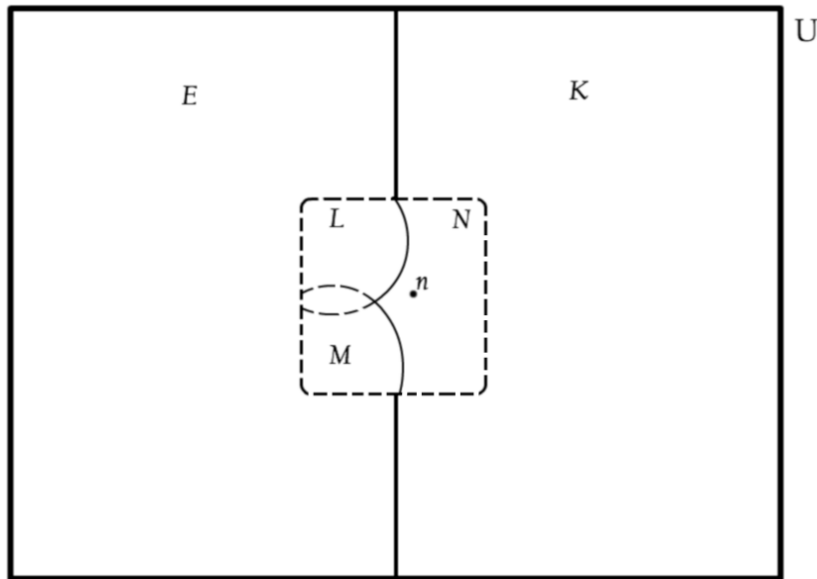
$$\therefore n \in K$$

And we can also model the analysis graphically (representing only those features of the universe of discourse that are of immediate interest):

Note an interesting implication of this formalization: as a matter of logical structure, the Trump Card is equivalent to a special case of the Tertium Quid in which $N = (A \cap H) = Z$.

¹⁸⁸ See *supra* note 186 and accompanying text.

Figure 21



One interesting feature of the *Tertium Quid* is its durability in a common law system relative to the *No True Scotsman*. Since it was handed down, the framework of *Rogers* has been adopted by a growing number of federal courts hearing cases in which a trademark claim runs into a First Amendment defense. The Ninth Circuit adopted *Rogers*' approach in *Mattel, Inc. v. MCA Records, Inc.*,¹⁸⁹ and applied it to the content (as opposed to merely the title) of expressive works in *E.S.S. Entertainment 2000, Inc. v. Rock Star Videos, Inc.*¹⁹⁰ (The Second Circuit itself had extended *Rogers* to the content of expressive works almost immediately.¹⁹¹) Other Circuit Courts of Appeal have similarly adopted

¹⁸⁹ 296 F.3d 894 (9th Cir. 2002).

¹⁹⁰ 547 F.3d 1095 (9th Cir. 2008).

¹⁹¹ *Cliffs Notes, Inc. v. Bantam Doubleday Dell Publ'g Grp., Inc.*, 886 F.2d 490 (2d Cir. 1989).

Rogers in one form or another.¹⁹² As of this writing it is the most cited American case dealing with the expressive use of trademarks.¹⁹³

As noted above,¹⁹⁴ the recursive nature of common-law adjudication means that the use of a No True Scotsman in an existing area of doctrine always has the potential to establish and propagate a new intensional definition. The Tertium Quid, however, provides such a definition, and in so doing has the potential to bring considerable stability to an area of doctrine where consistency is lacking. That is, a *precedent* that announces a Tertium Quid as the rationale for its decision purports to instruct *future* courts that in the case of this particular doctrinal contradiction, they ought to resolve the contradiction in a particular way. The Trump Card similarly can stabilize inconsistent doctrine when announced as the rationale for a decision: It instructs future courts to apply one conflicting rule and ignore the other.¹⁹⁵ The No True Scotsman lacks this quality, and does so because of its *logical form*: It speaks to the membership relations between a case and a legal set, but says nothing about the relationships between the legal sets whose inconsistency required its invocation. In the next Part, we will attempt to connect this relationship between the *form* of the various contradiction-avoidance strategies and the evolution of the universes of discourse in which they arise to deeper theoretical issues in the analysis of common-law legal systems.

¹⁹² See, e.g., *Radiance Found., Inc. v. N.A.A.C.P.*, 786 F.3d 316, 329 (4th Cir. 2015); *Univ. of Alabama Bd. of Trustees v. New Life Art, Inc.*, 683 F.3d 1266, 1278 (11th Cir. 2012); *Parks v. LaFace Records*, 329 F.3d 437, 451 (6th Cir. 2003); *Cardtoons, L.C. v. Major League Baseball Players Ass'n*, 95 F.3d 959, 971–72 (10th Cir. 1996); *Westchester Media v. PRL USA Holdings, Inc.*, 214 F.3d 658, 665 (5th Cir. 2000); cf. *Hart v. Elec. Arts, Inc.*, 717 F.3d 141, 157 (3d Cir. 2013) (“While the [*Rogers*] Test may have a use in trademark-like right of publicity cases, it is inapposite here.”).

¹⁹³ Westlaw Key Number Search 382Tk1524 (Trademarks>Violations of Rights>Defenses, Excuses, and Justifications>Justified or Permissible Uses>Expressive Use; Commentary) (last visited Mar. 25, 2019); see also Westlaw Key Number Search 92k1604 (Constitutional Law>Freedom of Speech, Expression, and Press>Trade or Business>Trademarks and Trade Names) (last visited Mar. 25, 2019) (second most cited case, after *Friedman v. Rogers*, 440 U.S. 1 (1979), a commercial speech/occupational licensing case).

¹⁹⁴ See *supra* notes 172–74 and accompanying text.

¹⁹⁵ Again, this makes sense insofar as we have established that the Trump Card is formally a special case of the Tertium Quid. See *supra* note 187.

E. *Summary*

We are now in a position to provide a graphical summary and decision tree for the various formal contradiction-resolution strategies reviewed in this Part. Both are provided in Figures 22 and 23 below. This summary presents an array of formal options by which an advocate may argue—or a judge may decide—an issue in an area of law where rules conflict. The next Part will discuss some implications of the model heretofore developed, as well as some of its limitations and its interactions with other aspects of legal theory.

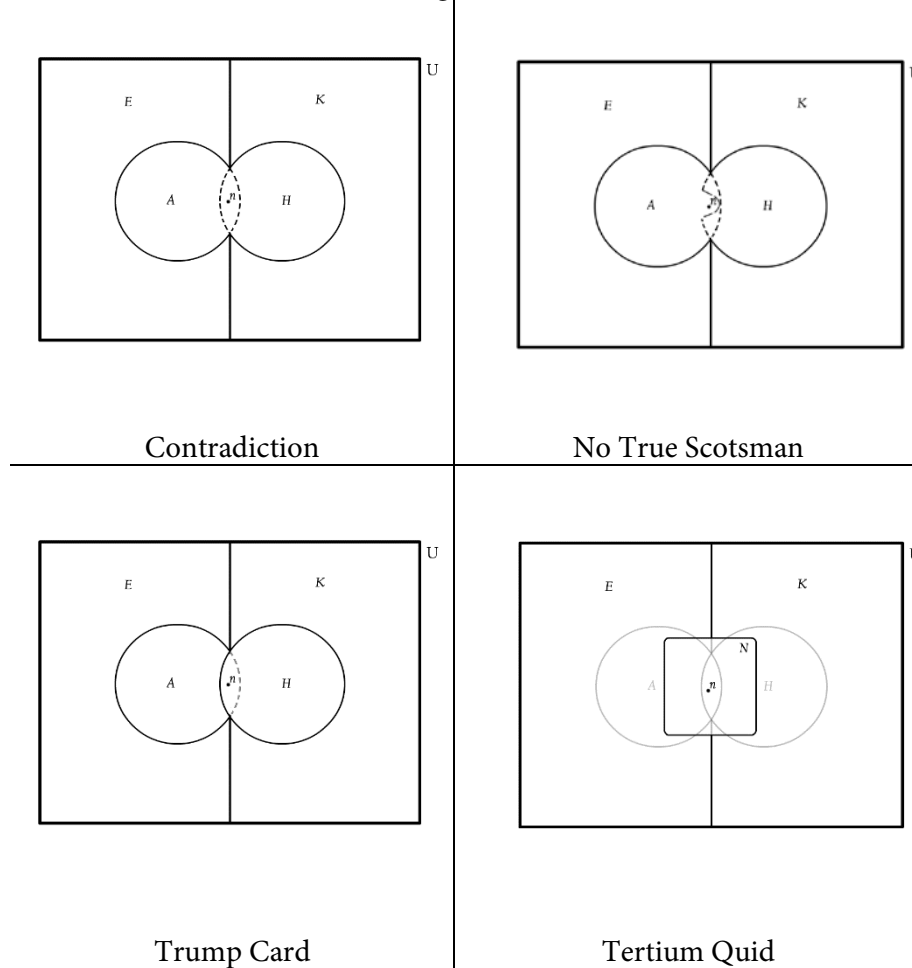
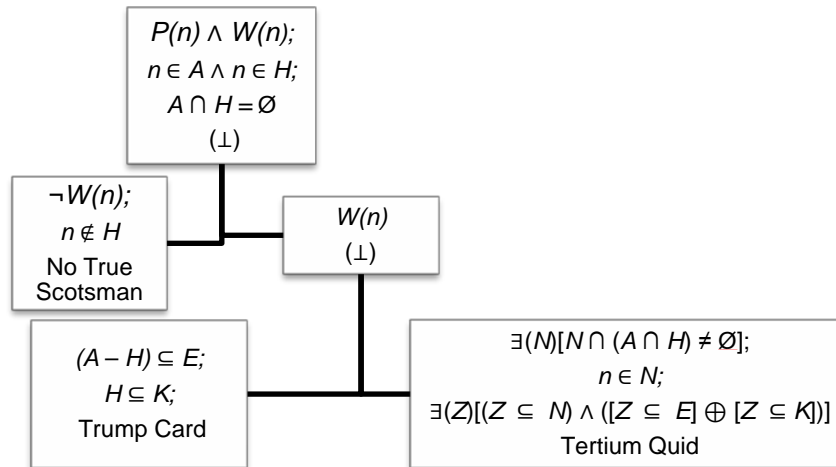
Figure 22

Figure 23



IV. IMPLICATIONS AND LIMITATIONS

With a thorough understanding of how various moves in legal practice can be modeled using set theory, we can begin to explore some of the implications of these models, and the ways in which a set-theoretic model of legal practice interfaces with other aspects of legal theory. In this Part, I will sketch out ways in which the set-theoretic model developed above can provide a framework within which more established approaches to legal theory can do their work. The model does so principally by cutting through ambiguities of language that bedevil both the practice of law and the arguments of legal theorists. In doing so, it provides reasons to believe that certain topics of jurisprudential interest and debate can be helpfully understood as emergent—and perhaps inevitable—properties of the set-theoretic structure of a common-law system.

A. *What the Set-Theoretic Model Is, and What It Is Not*

I have been referring to the set-theoretic model of legal practice as a *formal* model to distinguish it from *substantive* models of law or legal practice—the typical concerns of legal theory. Legal theorists generally deal with what we might call the *substance* of legal systems—ontological

concepts such as Hart's "rule of recognition"¹⁹⁶ or Dworkin's "grounds of law"¹⁹⁷; prescriptive agendas that motivate debates between textualists and purposivists, or originalists and dynamists;¹⁹⁸ psychological theories of what judges "actually" do, such as those that divide realists and anti-realists;¹⁹⁹ and so on. The set-theoretic model developed in this Article can accommodate any of these substantive theories, because its goals are relatively modest in comparison to such sweeping ideological projects. It aims only to provide a formal language to describe the *behaviors* that put any such substance into practice. As I hope will become apparent, this modest ambition is still a worthy one, and illuminates important aspects of legal practice and even of legal theory, but it necessarily omits certain features that some theorists might deem important to a model of a legal system.

First, the set-theoretic model lacks any theory of empirical validation. It cannot tell us when a particular *fact in the world* is true or false. Thus, when evaluation of a case's membership in a legal set requires some knowledge about the state of the world, we require some means from outside the model for obtaining and evaluating that knowledge. These external means may also be channeled through law, and that body of law may itself be modeled using sets—as with the law of evidence, rules allocating fact-finding responsibility, burdens of proof and persuasion, and procedural law generally.²⁰⁰ But a set-theoretic model of substantive law takes the outputs of these aspects of the legal system as inputs into other universes of discourse; they are complementary parts of a larger whole. The key point is that the set-theoretic model itself does not purport to provide any mechanism for determining such empirical facts—though it may be able to model systems (particularly legal systems) that do.

¹⁹⁶ HART, *supra* note 8, at 94–110.

¹⁹⁷ RONALD DWORKIN, LAW'S EMPIRE 4–6 (1986).

¹⁹⁸ Even the recent empirical turn in the debate is substantive, to the extent it seeks to discern what motivates judicial acts of interpretation. See *generally* William Baude, *Is Originalism Our Law?*, 115 COLUM. L. REV. 2349 (2015).

¹⁹⁹ See *generally* Leiter, *supra* note 12.

²⁰⁰ An especially persuasive account of how legal systems arrive at such determinations can be found in the "fuzzy logic" model developed by Kevin Clermont. Kevin M. Clermont, *Conjunction of Evidence and Multivalent Logic*, in LAW AND THE NEW LOGICS (H. Patrick Glenn & Lionel D. Smith eds., 2017).

Second, the set-theoretic model lacks what Schauer calls a “rule of relevance”: an “organizing standard specifying which similarities [between cases] are important and which we can safely ignore.”²⁰¹ Such rules are a focus of other efforts to model the logic of law and other deontic systems,²⁰² and their omission is felt in two aspects of my model. First, knowing how to model the formulation and conflict of legal rules, or even the resolution of that conflict, does not necessarily help us identify whether a particular rule is implicated in a particular case. It does not, for example, provide any route to *generate* the statement “this series of events may give rise to a negligence claim,” or “this statute may pose a free exercise problem,” or “this litigant may have a justification defense.” Second, understanding that legal sets may be constructed extensionally from collections of cases does not tell us (beyond identification of a prevailing party) how to identify which cases belong together and which ones do not, let alone how best to formulate an intensional predicate for a set constructed in this way. For either of these important tasks, we need some theory of relevance from outside the model. In short, the model cannot tell us *which* set predicates to evaluate against a particular case or group of cases; it can only provide a framework for that evaluation once the relevant predicates have been identified (or constructed).²⁰³ For the present, I am content to hypothesize that such rules of relevance are, like the conflict-resolution strategies described above, best understood as emerging from legal *practice* rather than supposedly abstract legal substance.

²⁰¹ Schauer, *supra* note 98, at 577–78.

²⁰² See, e.g., NAVARRO & RODRIGUEZ, *supra* note 7, at 166–75 (citing CARLOS ALCHOURRÓN & EUGENIO BULYGIN, *NORMATIVE SYSTEMS* 103, 107 (1971)).

²⁰³ These missing theories of relevance are likely the province of professional training, and more generally of social context. See Schauer, *supra* note 98, at 578 (rules of relevance “are contingent upon both time and culture”); Pierre J. Schlag, *Rules and Standards*, 33 *UCLA L. REV.* 379, 407 (1985) (“After all, the question whether a piece of text (such as a rule or a standard) applies to a given context is a function of context.”); Jack M. Balkin, *The Framework Model and Constitutional Interpretation*, in *PHILOSOPHICAL FOUNDATIONS OF CONSTITUTIONAL LAW* 241, 248 (David Dyzenhaus & Malcom Thorburn eds., 2016) (“[A]t any point in time, some proposed interpretations are ‘off-the-wall,’ while others are plausible or ‘on-the-wall,’ even if they are not necessarily the best interpretation. . . . [T]he boundary between what people regard as reasonable and unreasonable is not fixed; it can change as a result of legal discussion and political mobilization.”).

Third, and finally, the model does not itself provide any guidance on which of the three contradiction-avoidance strategies a legal actor *ought* to use in any particular situation. We could imagine that such guidance might be part of a legal system, as a matter of positive law. We have already noted examples of such guidance in our own legal system: the Model Penal Code's treatment of lesser included offenses seems to require the Trump Card;²⁰⁴ *Kremen's* preference for extensional set-building implies that inconsistent cases should be resolved via the Tertium Quid.²⁰⁵ We can call these types of purported constraints on the set-theoretic moves available to legal actors *reconciliation rules*. A reconciliation rule, in this view, is any aspect of a legal system that purports to compel an actor within that system to choose one contradiction-avoidance strategy rather than another. As I will argue below, one of the most important implications of a set-theoretic model is that it casts serious doubt on the possibility that such reconciliation rules can *in themselves* be effective in constraining the behavior of legal actors.

Even with these limitations, the types of analyses reviewed in this Article can be extremely helpful in organizing one's thinking about how lawyers and judges do what they do (as I hope the previous Parts of this Article have demonstrated). Moreover, they can be helpful in clearing up ambiguities in theoretical discussions about law. That is not because the set-theoretic model is itself a model of what law "is" in some metaphysical sense, or that it is a prescriptive model of what judges or lawyers ought to do, or even that it describes what they believe they do or ought to do. Rather, it is because the set-theoretic model provides a language and a set of tools that allows the thinkers who investigate such problems to avoid talking past one another, and exposes some features of *law as a social practice* that emerge from the structure of the legal system itself. The remainder of this Part will sketch a few examples.

²⁰⁴ See *supra* Section II.B.2.

²⁰⁵ See *supra* Section II.C.

B. *Precision and Generality: Rules and Standards, Policies and Principles*

The rules-versus-standards dichotomy is one of the most familiar in legal theory. It has ossified into a routine set of pro-and-con arguments regarding the merits of bright-line rules and flexible standards: a “dialectic” that Pierre Schlag despairingly (or cynically) characterized as “irreducible.”²⁰⁶ Duncan Kennedy famously attempted to tie the distinction to ideological and political strands in adjudication: mapping “form” (the rule/standard design choice) to “substance” (the political and ideological implications of the design choice).²⁰⁷ The late Justice Scalia identified rules with the constraint of judicial discretion, and with his favored interpretive theory of originalism.²⁰⁸ Louis Kaplow built an economic model around the dichotomy, tying it to gradients in information costs and administration costs.²⁰⁹ In these scholarly treatments, the difference between rules and standards is treated as if it is usefully reducible to a single dimension or spectrum. But the logical characteristics that distinguish what we call rules from what we call standards are more complex than a one-dimensional gradient. And the failure to properly distinguish the multiple *formal* distinctions among what legal theorists refer to as “rules” or “standards” results in the jumbling together of distinct theoretical issues. The set-theoretic model can help cut through this ambiguity.

One sense in which we might distinguish a rule from a standard involves evaluation of membership relations. In particular, it involves the extent to which a set’s boundary is clearly delineated by an intensional predicate. By formulating legal directives as sets, and diagramming them with Venn Diagrams, I do not mean to suggest that such directives always draw sharp boundaries between close cases. Whether the intensional predicate of a doctrinally relevant set clearly

²⁰⁶ Schlag, *supra* note 203, at 426.

²⁰⁷ See generally Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685 (1976).

²⁰⁸ See generally Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175 (1989).

²⁰⁹ See generally Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992).

identifies what is in and what is out of that set depends on the words in which that predicate is framed and the state of the world (or our knowledge of the state of the world) against which it is evaluated. And when predicates are specified in natural language—as in a legal system they generally are—there will often be uncertainty over whether they are a truthful description of a particular state of the world, just because the correspondence of words to facts in the world is not perfect.²¹⁰

We can understand this aspect of the rules/standards dichotomy to refer to the *precision* of legal directives.²¹¹ Precision, as used here, refers to the degree of correspondence between natural language and the world: A more precise predicate *reduces uncertainty in evaluating the membership relations* between that predicate and facts in the world. For example, the directive “The terms of the President and Vice President shall end at noon on the 20th day of January”²¹² more precisely identifies constraints on the office of President than the directive “he shall take Care that the Laws be faithfully executed”²¹³; the classic precise predicate is the prohibition that “no person shall drive a vehicle at a speed in excess of fifty-five miles per hour.”²¹⁴

But there is another very different sense in which we might distinguish a rule from a standard. In the previous Part we focused on the possibility that two distinct legal directives might conflict, but of course it is also possible that they could coincide. That is: Two different legal directives might both require the same result in a particular case, but for different reasons. Often when there are two legal directives that coincide in this way, it is because one of the directives is framed in such a way as to address only a subset of the cases governed by the other directive. For example, the Equal Protection Clause—“nor shall any State . . . deny to any person within its jurisdiction the equal protection

²¹⁰ See generally Schauer, *supra* note 98.

²¹¹ This is the sense in which Timothy Endicott evaluates the “precision” (or its converse, “vagueness”) of legal texts, and it may be decomposed further into imprecision along various semantic dimensions, none of which are important to the distinction I am attempting to draw. See generally Timothy Endicott, *The Value of Vagueness*, in PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW 14 (Andrei Marmor & Scott Soames eds., 2011).

²¹² U.S. CONST. amend. XX, § 1.

²¹³ U.S. CONST. art. II § 3.

²¹⁴ N.Y. VEH. & TRAF. LAW § 1180(b) (2016).

of the laws”²¹⁵—has purchase on legal disputes involving nearly every area of human interaction.²¹⁶ Conversely, the command of the Voting Rights Act of 1965—that “[n]o voting qualification or prerequisite to voting, or standard, practice, or procedure shall be imposed or applied by any State or political subdivision to deny or abridge the right of any citizen of the United States to vote on account of race or color”²¹⁷—deals only with the narrower domain of state electoral rules. For legal disputes within that narrower domain, the Voting Rights Act might well generate the same outcomes as the Equal Protection Clause, even though the statute says nothing about other areas—such as education, or housing, or government benefits—that might come within the command of the constitutional provision.

We can give this difference in scope a name: call it *generality*. Generality is a measure of the proportion of the conceptual space within a relevant universe of discourse to which a particular predicate may be relevant. In a nutshell, within our set-theoretic model, *generality* is a measure of the size of a legal set; *precision* is a measure of the sharpness of its boundary.²¹⁸ Both precision and generality are relevant to the characterization of a legal directive as either a rule or a standard, but legal theorists often blur the two dimensions together.²¹⁹ The result is that they conflate two very different points: one about the nature of legal

²¹⁵ U.S. CONST. amend. XIV, § 1.

²¹⁶ Gerald Gunther, *Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1, 8–9 (1972).

²¹⁷ Voting Rights Act of 1965, Pub. L. No. 89-110, § 2, 79 Stat. 437 (1965).

²¹⁸ Cf. Adam M. Samaha, *Levels of Generality, Constitutional Comedy, and Legal Design*, 2013 U. ILL. L. REV. 1733, 1744, 1749 (2013) (“[A]bstractness includes vagueness and other hindrances to ease of understanding. . . . Breadth . . . [involves the dimension along which a] source might be understood to bear on a broad range of circumstances or only a narrow corner of social life.”) (emphasis omitted).

²¹⁹ One notable exception is Cass Sunstein’s theory of judicial minimalism, which—though it does not directly engage the rules/standards dichotomy—reflects in its concepts of “depth” and “width” features analogous (though not identical) to the ideas of “precision” and “generality” that I describe here. See CASS R. SUNSTEIN, *ONE CASE AT A TIME: JUDICIAL MINIMALISM ON THE SUPREME COURT* 10–19 (1999). For example, Sunstein’s concept of “depth” has to do with the depth of theoretical principle underlying a particular disposition; generality does not speak to substantive theory, but rather to the concepts of inclusion—the idea that a particular case is representative of a broader set of cases, and that those cases can be tied together by a sufficiently encompassing intensional definition. See *id.*

directives framed in natural language, and the other about the behavior of legal actors engaging with one another over time.

Take this representative passage from Kathleen Sullivan's Foreword to the 1992 Supreme Court Issue of the Harvard Law Review:

A legal directive is "rule"-like when it binds a decisionmaker to respond *in a determinate way* to the presence of *delimited triggering facts*. Rules aim to confine the decisionmaker to facts, leaving irreducibly arbitrary and subjective value choices to be worked out elsewhere. . . . A legal directive is "standard"-like when it tends to collapse decisionmaking back into the direct application of the *background principle or policy* to a fact situation. Standards allow for the decrease of errors of under- and over-inclusiveness by *giving the decisionmaker more discretion* than do rules. Standards allow the decisionmaker to take into account all relevant factors or the totality of the circumstances. *Thus, the application of a standard in one case ties the decisionmaker's hand in the next case less than does a rule.*²²⁰

Dean Sullivan's initial definition of a rule invokes the dimension of precision—favoring "delimited" predicates and avoiding "arbitrary and subjective" ones. But her definition of a standard invokes the dimension of generality: a "background principle or policy" rather than the narrower, more focused directives that might be included within, or consistent with, such a policy. Moreover, her discussion of rule-like directives is about the *characteristics of rules* as tools for constraining legal actors. Rules (or standards) are the subject, legal actors the object. But the last sentence slides into the passive voice, masking a reversal of roles. When she discusses "the application . . . in one case" of a rule or a standard, the subject is *the legal actor* who does the applying; the rule (or standard) has become the object. That is, she has shifted into discussion of *the selection of tools by legal actors*. These two pairs of conflated concepts—precision with generality, and characteristics of rules with characteristics of behavior—bear unpacking.

²²⁰ Kathleen M. Sullivan, *The Supreme Court, 1991 Term—Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 58–59 (1992) (emphasis added).

The problem with blurring together the concept of precision with the concept of generality is that the two are not strictly correlated.²²¹ To be sure, the more general a legal directive is, the less precise it is likely to be, if only because a directive that purports to apply to a broad and diverse range of states of the world will necessarily be framed in words that have purchase on multiple diverse aspects of the world.²²² But one can imagine highly general directives that are quite precise—for example: “always decide a legal dispute in favor of the party with the higher adjusted gross income on their previous year’s income tax return.” And of course, an imprecise directive might be quite specific—for example: “No person shall open the door of a motor vehicle on the side available to moving traffic unless and until it is reasonably safe to do so.”²²³ Perhaps the most familiar example of the mutual independence of precision and generality can be seen in legislative delegations of regulatory authority. For example, Section 619 of the Dodd-Frank Act (the so-called “Volcker Rule”) prohibits banks from “acquiring or taking positions in [certain financial instruments] principally for the purpose of selling in the near term”;²²⁴ federal agencies then promulgated a presumption that a position held for less than sixty days constituted a violation of the rule.²²⁵ The generality of both directives is the same—they are directed to the same category of human conduct—but the precision of the directives is different—a presumption based on the passing of a fixed period of time as compared with a description of the “purpose” of the directive’s target.

The distinction between the precision and generality of legal directives has important implications for the second set of concepts conflated in Dean Sullivan’s discussion: the characteristics of legal

²²¹ See Samaha, *supra* note 218, at 1749–50 (diagramming the interaction between the author’s posited dimensions of “abstractness” and “breadth,” which have affinities with precision and generality as defined herein).

²²² Cf. Laurence H. Tribe & Michael C. Dorf, *Levels of Generality in the Definition of Rights*, 57 U. CHI. L. REV. 1057, 1058 (1990) (“[A]t what level of generality should the Court describe the right previously protected and the right currently claimed? The more abstractly one states the already-protected right, the more likely it becomes that the claimed right will fall within its protection.”) (emphasis omitted).

²²³ N.Y. VEH. & TRAF. LAW § 1214 (1964).

²²⁴ 12 U.S.C. §§ 1851(a)(1)(A), (h)(4), (h)(6) (2018).

²²⁵ 17 C.F.R. § 255.3(b)(2) (2014).

directives and the practices of legal actors. This is because of the possibility of multiple directives at different levels of generality bearing on any given case, and the need for legal actors to choose from among them. These choices imply a dimension of legal discretion that goes beyond the imprecision of natural language—and the concomitant indeterminacy of the application of any particular rule to a case—to the underlying structure of a common-law system. Put simply, legal actors not only decide whether their case falls within the language of a legal directive, but also which of multiple directives at different levels of generality they will use to evaluate the case. Moreover, these two dimensions of discretion are independent of one another, and the latter dimension depends not (or not only) on the language in which a directive is framed, but—as the rest of this Part will explain—on the interactions among legal actors over time.

The set-theoretic model is obviously not *necessary* for an appreciation of these distinctions lying under the surface of the rules/standards dichotomy. But, I contend, it is extremely *useful* in this regard, precisely because such distinctions emerge effortlessly from its foundational distinction between membership and inclusion, mapping to the relationship between cases and legal directives on the one hand, and among multiple legal directives on the other. The model thus not only brings clarity and precision to the debate over the rules/standards dichotomy, it also foregrounds (helpfully, in my view) the role of legal actors and their behaviors in the dilemmas that dichotomy supposedly presents.

C. *Dimensions of Discretion: Precedent*

Another example from intellectual property law demonstrates the utility of the set-theoretic model in understanding how theoretical concerns about law emerge from the structure of interactions among legal actors over time. In this case, the theoretical concern is the nature of precedent, and particularly its ability to constrain legal actors. The doctrine of aesthetic functionality in trademark law, as explained in the Supreme Court case of *Qualitex Co. v. Jacobson Products Co.*²²⁶ and its

²²⁶ 514 U.S. 159 (1995).

progeny, prohibits any firm from attempting to protect any product feature as a trademark “if exclusive use of the feature would put competitors at a significant non-reputation-related disadvantage”²²⁷ by allowing monopolization of features that are “intended not to identify the source [of the product], but to render the product itself more useful or more appealing.”²²⁸ However, trademark law does allow producers to claim exclusive rights to “trade dress” product features²²⁹ provided that they have established “secondary meaning” for such features—that is, if a substantial portion of the consuming public has come to associate the product feature with one particular producer.²³⁰ A doctrinal conflict therefore arises with respect to product features that consumers demand for *both* reputation-related *and* non-reputation-related reasons. In recent litigation over Christian Louboutin’s red-soled stiletto shoes, the record presented exactly such a conflict. Louboutin himself had admitted that the red soles of his shoes were desirable because they were “engaging, flirtatious, memorable and . . . sexy,” but the record also established that “in the high-stakes commercial markets and social circles in which these things matter a great deal, the red outsole became closely associated with Louboutin.”²³¹

In deciding Louboutin’s motion for a preliminary injunction against Yves Saint Laurent (which was marketing shoes that were red all over, including their soles), District Judge Marrero invoked the Tertium

²²⁷ *Id.* at 165; *see also* *Traffix Devices v. Mktg. Displays*, 532 U.S. 23, 33 (2001) (“It is proper to inquire into a ‘significant non-reputation-related disadvantage’ in cases of aesthetic functionality.”).

²²⁸ *Wal-Mart Stores v. Samara Bros.*, 529 U.S. 205, 213 (2000).

²²⁹ Trade dress “involves the total image of a product and may include features such as size, shape, color or color combinations, texture, graphics, or even particular sales techniques.” *Two Pesos, Inc. v. Taco Cabana, Inc.*, 505 U.S. 763, 764 n.1 (1992) (quoting *John H. Harland Co. v. Clarke Checks, Inc.*, 711 F.2d 966, 980 (11th Cir. 1983)). Trade dress may include not only aspects of a product’s packaging but of its design as well.

²³⁰ *Wal-Mart Stores*, 529 U.S. at 211 (quoting *Inwood Labs., Inc. v. Ives Labs., Inc.*, 456 U.S. 844, 851 & n.11 (1982)) (“[A] mark has acquired distinctiveness, even if it is not inherently distinctive, if it has developed secondary meaning, which occurs when, ‘in the minds of the public, the primary significance of a [mark] is to identify the source of the product rather than the product itself.’”) (modification in original).

²³¹ *Christian Louboutin S.A. v. Yves Saint Laurent Am., Inc.*, 778 F. Supp. 2d 445, 447–48 (S.D.N.Y. 2011).

Quid by situating the case as a member of a newly identified set of cases involving the use of particular colors in fashion goods:

[T]he Court must decide whether there is something unique about the fashion world that militates against extending trademark protection to a single color, although such registrations have sometimes been upheld in other industries. . . . [T]he Court cannot conceive that the Lanham Act could serve as the source of the broad spectrum of absurdities that would follow recognition of a trademark for the use of a single color for fashion items.²³²

On appeal, the Second Circuit did not disturb Judge Marrero's *disposition* of the preliminary injunction motion, but it emphatically rejected his inclusion-based Tertium Quid move, instead opining that he should have simply evaluated the membership relation between the case and a legal directive extracted from the text of the *Qualitex* opinion:

Qualitex requires an individualized, fact-based inquiry into the nature of the trademark, and cannot be read to sanction an industry-based *per se* rule. The District Court created just such a rule. . . . Even if *Qualitex* could be read to permit an industry-specific *per se* rule of functionality (a reading we think doubtful), such a rule would be neither necessary nor appropriate here.²³³

However, rather than undertaking such an inquiry, the Second Circuit disposed of the case by evaluating a *different* membership relation. That is, rather than finding that the claimed mark was not functional under the test of *Qualitex*—as it said the District Court should have done—the Second Circuit concluded the trademark as claimed was not source-identifying.²³⁴ This avoided the contradiction not only by using a

²³² *Id.* at 451, 457.

²³³ *Christian Louboutin S.A. v. Yves Saint Laurent Am. Holding, Inc.*, 696 F.3d 206, 223 (2d Cir. 2012).

²³⁴ The Second Circuit opinion also determined—in what can only be categorized as an advisory opinion given its other holdings—that Louboutin's claimed trademark was only partially (as opposed to completely) invalid. *See id.* at 227–28 (“[T]he record fails to demonstrate that the secondary meaning of the Red Sole Mark extends to uses in which the sole does not contrast with the upper—in other words, when a red sole is used on a monochromatic red shoe. . . . We therefore instruct the Director of the Patent and Trade Office to limit the registration of the Red Sole Mark to only those situations in which the red lacquered outsole contrasts in color with the adjoining ‘upper’ of the shoe. . . . [W]e hold that the Red Sole Mark is

different logical strategy, but by resort to different rules within the universe of discourse.

The difference between the District Court and Court of Appeals opinions in *Louboutin v. Yves Saint Laurent* has little to do with the uncertain application of legal rules to the specific dispute before the courts. The *precision* of the doctrine of aesthetic functionality simply played no role. In both opinions, Louboutin's claimed trademark was found unenforceable against Yves Saint Laurent—in the District Court because it was functional, in the Second Circuit because it was not source-identifying. Neither rationale is *substantively* inconsistent with the other. The primary reason to reverse the District Court's resolution of the doctrinal conflict it faced was *formal*, and hinged on the recursive structure of a common-law system. That is, the Second Circuit's reversal changed the *prospective* effect of the disposition of the dispute on the structure of the universe of discourse in which the conflict arose. The District Court would have reshaped that structure to prevent the contradiction it encountered from arising in the future *as a matter of logical form*. The Court of Appeals preferred to leave the existing structure—and its potential for contradiction—intact, rather than alter the model with a novel intensional definition of a heretofore unrecognized set (i.e., “color as a trademark for fashion items”).

The *Louboutin* case demonstrates how courts can and do shape their behaviors with a view not only to the substance of their decisions, but to their effect on the formal structure of the universe of discourse within which those decisions are situated. It is a reminder of the theoretical insight that precedent is not only “backward-looking”—insofar as it purports to base a present adjudication on adjudications of the past—but also “forward-looking”—insofar as the *reasons* for today's decision purport to bind future legal actors.²³⁵ The legal directives announced in judicial opinions—like any legal authorities—may be framed in language that is more or less precise. But they may also be pitched at different levels of generality. Thus, on their face, at least,

valid and enforceable as modified. . . Having limited the Red Sole Mark as described above, and having established that the red sole used by YSL is not a use of the Red Sole Mark, it is axiomatic that *we need not—and should not—address . . . whether the modified Mark is functional.*”) (emphasis added).

²³⁵ DWORKIN, *supra* note 197, at 225; see generally Schauer, *supra* note 98.

inclusion moves—like the Tertium Quid and the Trump Card—make more substantial claims on the behavior of future legal actors than membership moves—like the No True Scotsman. For example, as a forward-looking precedent, the opinion: “we cannot find substantial equality in the educational opportunities offered white and Negro law students by the State [of Texas]”²³⁶ makes fewer claims on future legal disputes than the opinion: “[s]eparate educational facilities are inherently unequal.”²³⁷

Legal actors with sophisticated understandings of the types of logical moves available to them can deploy those moves strategically, in a bid to shape the universe of discourse in a preferred direction over time.²³⁸ The normative desirability of such strategic behavior is beyond the scope of this article. But from a purely descriptive standpoint, we are confronted with two questions central to the nature of precedent: First, whether we think such strategic efforts are likely to be successful in influencing the behavior of future legal actors. And second, whether law can constrain such strategic efforts. The set-theoretic model does not provide complete answers to these questions, but it suggests that both answers are interconnected. Specifically, the indeterminacy baked into the *structure* of law as a social practice limits the power of any particular legal actor to leverage that indeterminacy in results-oriented ways.

D. *Meta-Rules: Sets All the Way Down*

How might a legal actor constrain, in advance, a different legal actor’s later evaluation of a case? Such a project is beset with practical problems, which Hart famously identified with epistemic deficiencies of the would-be constrainer: “ignorance of fact” and “indeterminacy of aim.”²³⁹ But it also raises theoretical problems arising from the structure of law as a social practice, corresponding to the dimension of precision,

²³⁶ *Sweatt v. Painter*, 339 U.S. 629, 633 (1950).

²³⁷ *Brown v. Bd. of Educ.*, 347 U.S. 483, 495 (1954).

²³⁸ See generally Oona A. Hathaway, *Path Dependence in the Law: The Course and Pattern of Legal Change in a Common Law System*, 86 IOWA L. REV. 601 (2001) (arguing that precedent can create a kind of “path dependence” that constrains future doctrinal development).

²³⁹ See HART, *supra* note 8, at 127–28; see also H. L. A. Hart, *Discretion*, 127 HARV. L. REV. 652, 661–63 (2013).

the dimension of generality, and the set-theoretic interaction of directives that vary along both dimensions as precedents making claims on the behavior of legal actors over time.

The ineliminable discretion that results from the imprecision of the language in which legal directives are wrought is a well-worn theoretical chestnut. For example, Hart's famous example of a legal rule forbidding vehicles in the park²⁴⁰ is generally used as the jumping-off point for debates over the relation of a number of hypothetical states of the world to the word "vehicle"—that is, about the proper way to ascertain the relationship between a legal text and a state of the world.²⁴¹ The theoretical architecture of interpretation—of which Hart's exercise is a classic exemplar—is largely concerned with the application of rules to states of the world and with minimizing or resolving the uncertainty of such an exercise. So, for example, when originalists argue (either amongst themselves or with adversaries such as living constitutionalists) over the proper interpretation of legal texts,²⁴² they are generally talking about interpretation of the meaning of the words of those texts—in light of whatever sources of meaning they deem acceptable—and whether that meaning is descriptive of the state of the world the legal actor is trying to evaluate.²⁴³

²⁴⁰ H. L. A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593, 607 (1958).

²⁴¹ See generally Pierre Schlag, *No Vehicles in the Park*, 23 SEATTLE U. L. REV. 381 (1999). Typically the text in question is a statute or a constitutional provision, but the same general approach applies to rules derived from the text of judicial opinions. See Schauer, *supra* note 98, at 580–81 (“[T]he articulated characterization [within a judicial opinion] acts like a specifically formulated rule. . . . Where there is an articulated characterization, therefore, the question whether precedent can constrain may collapse into the question whether rules can constrain.”); see also Scalia, *supra* note 208, at 1177 (“[B]y making the[ir] mode of analysis relatively principled or relatively fact-specific, the courts can either establish general rules or leave ample discretion for the future.”).

²⁴² For a recent brief overview of the debate over originalism—in all its flavors—see Baude, *supra* note 198, at 2351–63 (2015). For a review of the distinctions between originalists and a leading competing school—living constitutionalism—see generally Ethan J. Leib, *The Perpetual Anxiety of Living Constitutionalism*, 24 CONST. COMMENT. 353 (2007).

²⁴³ This is true even for those more recent and subtle theories that draw a distinction between “semantic meaning” and “legal meaning,” such as the interpretation/construction distinction championed by Larry Solum. See Lawrence B. Solum, *The Interpretation-Construction Distinction*, 27 CONST. COMMENT. 95, 103–08 (2010).

We have come to identify such an exercise with the set-theoretic concept of membership, and the uncertainty inherent in membership evaluations with the imprecision of natural language. But let us assume that a legal actor concludes that a particular directive is not predicative of the case they must resolve—perhaps through a contestable exercise of discretion in the evaluation of the case with respect to the directive’s imprecise language. How then should the legal actor resolve the case to a binary outcome? The notion of a universe of discourse implies that the case must be a member of some *other* set, predicated by some *other* legal directive.

Perhaps this other directive is simply a default rule at a high level of generality—an example of the deontic-logical concept of “defeasibility.”²⁴⁴ Burdens of proof have this character: they direct an outcome in favor of the non-burdened party whenever the burdened party fails to establish membership in some set that would direct an outcome in their favor. But not all alternative sets have this default (or complementary) relationship to sets predicated by less general legal rules. As we have seen, sometimes multiple legal rules are implicated in a particular case, and they may conflict or coincide. Thus, evaluating a case’s membership in a particular rule-based set is not the sum total of a legal actor’s work. Rather, the actor must *construct an internally consistent universe of discourse* comprised of whatever rule-based sets she believes are implicated by her case—each of them mapping to adjudicative outcomes—and then *situate her case within that universe of discourse*.²⁴⁵

This task is complicated by the dimension of generality, and the need to mind relationships among multiple legal directives pitched at different levels of generality. Nowhere is this complication more

²⁴⁴ See sources cited *supra* note 52 and accompanying text.

²⁴⁵ One very recent and quite helpful effort to model this type of exercise using (nonmonotonic) default logic is Lawsky, *supra* note 22. Lawsky’s use of defeasible logic to formally model complex statutory regimes involving exceptions and qualifications is a major advance in modeling a universe of discourse defined by a single statute, or even a group of related statutes such as the tax code. However, the problem of selecting which authorities, rules, or principles are implicated in a particular case, and how they ought to be interpreted against one another when they do not themselves say how they should be so interpreted, would remain even if we accept that the logic of the individual implicated directives is defeasible rather than monotonic.

apparent than in the institution of precedent. The discussion of *Kremen v. Cohen*²⁴⁶ illustrated that a precedent can be read either extensionally or intensionally. Extensionally, at the extreme, it may contribute nothing more to a universe of discourse than “the facts treated by the judge as material, and . . . his decision as based on them”²⁴⁷—an element of one of the complementary sets predicative of binary outcomes, rather than a set in its own right. Intensionally, at the opposite extreme, the “articulated characterization” offered by a court as a justification for its decision—its reasoned opinion—may operate similarly to legislation, providing a directive whose precision and generality may be arbitrarily determined by the opining court itself, and which “constrains the use of subsequent and inconsistent characterizations.”²⁴⁸ Between these extremes, lawyers and theorists may argue about the relative domains of holding and dicta, and about the similarity and difference between the facts of precedent and those of some instant case—in short, about the appropriate level of generality at which to frame some directive derived from the precedent.²⁴⁹ A single precedent thus encapsulates the potential for conflicts between multiple intensionally framed legal authorities (such as statutes or constitutional provisions).

Any effort to constrain a legal actor’s discretion will have to engage not only the actor’s evaluation of the relationship between the language of particular rules and the state of the world the actor is called on to evaluate according to such rules, but also her selection of *which* rules (at which level of generality) to include in her universe of discourse, as well as her selection of a strategy to resolve any conflicts among the rules she selects.²⁵⁰ The first task is the subject of rules of relevance, while the

²⁴⁶ See *supra* Section II.C.

²⁴⁷ Arthur L. Goodhart, *Determining the Ratio Decidendi of a Case*, 40 YALE L.J. 161, 182 (1930).

²⁴⁸ Schauer, *supra* note 98, at 579–80.

²⁴⁹ As Frederick Schauer has noted, every judicial opinion that contains more than a recitation of its facts and an outcome in some sense says more than is *strictly* necessary to the outcome of a case. Schauer, *supra* note 162, at 648 (“If a reason that can be narrower is for that reason dicta, then anything other than the announcement of an outcome is dicta.”). For a review of theoretical treatments of the holding/dicta dichotomy, see generally Michael Abramowicz & Maxwell Stearns, *Defining Dicta*, 57 STAN. L. REV. 953 (2005).

²⁵⁰ For an effort to formally reconcile this tension, see generally Horty, *supra* note 75.

second is the subject of reconciliation rules.²⁵¹ We can refer to such rules—which purport to direct a legal actor’s selection and application of doctrinal rules—as *meta-rules*. The question is whether an earlier-in-time legal actor—for example, a legislator enacting a statute or a judge announcing a precedent—can create such meta-rules so as to effectively constrain a later-in-time actor’s performance of these analytic tasks.

As I have said, the set-theoretic model, being purely descriptive, does not require or prescribe any such meta-rules. Nor will I defend any position on which such rules are desirable. But the model still allows us to draw some conclusions about their properties, to the extent they are included within a legal universe of discourse. To see why, we must consider that whatever the substance of these rules of relevance or reconciliation rules might be, *they are still rules*. That is, they purport to provide directives to guide the behavior of legal actors. They thus have all the set-theoretic characteristics of any other legal directive—including the dimensions of precision and generality. They can be modeled—and manipulated—using any of the set-theoretic relations and operations reviewed in this Article. They therefore may present similar—if not worse—problems of linguistic imprecision, and more importantly will similarly be subject to *formal* discretion on the part of legal actors who are, in turn, subject to them.

The famously contradictory judicial canons of statutory construction²⁵² provide a familiar example. One of Llewellyn’s most stark examples of “dueling canons” is his third: “Statutes are to be read in the light of the common law,” but “[t]he common law gives way to a statute which is inconsistent with it.”²⁵³ Both of these canons purport to do the work of reconciliation rules: to tell us how to resolve a conflict between contradictory substantive rules derived from statute and common law. Llewellyn’s point was that there is little guidance in the law of interpretation as to *which* of these contradictory rules should govern in any particular case—which we can recognize as a claim regarding the non-existence of useful or widely accepted rules of

²⁵¹ See *supra* Section IV.A.

²⁵² See generally Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons about How Statutes Are to Be Construed*, 3 VAND. L. REV. 395 (1949); Anita S. Krishnakumar, *Dueling Canons*, 65 DUKE L.J. 909 (2015).

²⁵³ Llewellyn, *supra* note 252, at 401.

relevance. Of course, even if this were not so, any such rule of relevance *would still be a rule* that would require evaluation by the same legal actor: The rule of relevance would merely push the problem of discretion to a different degree of generality. Moreover, the application of *either* canon identified by whatever rule of relevance we adopt would also necessarily be ambiguous as among the *formal* strategies we reviewed in Part III. For example, we do not know whether a common-law rule “giving way” requires an application of the Trump Card, or can be honored via application of a No True Scotsman or a Tertium Quid; if the last, we do not know what level of generality the new set should aim for. In short, neither canon informs a court how to determine whether a case before it ought instead to be evaluated against the canon *not followed*, nor whether the implicated statute (or common-law rule, as the case may be) should always, sometimes, or just in this instance prevail over the conflicting rule. If the answer to the latter question is “sometimes,” the canon provides no basis for determining when.

Thus, even if either of these canons provided definitive guidance on whether a statutory rule or a contradictory common-law rule ought to govern the outcome of a particular dispute as a matter of *relevance*, neither that rule nor the selected canon would provide guidance on the *form* such a disposition should take as a matter of *reconciliation*. In other words, even meta-rules—those rules that might purport to cabin *formal* discretion—are *themselves* subject to formal discretion. Whatever one’s *substantive* views regarding, for example, the centrality of text, the legitimate sources of legal meaning, the relevant historical or reflective posture for interpretation, the allocation of power between judges and legislators, the nature of *stare decisis*, or the distinction between holding and dicta—such substantive commitments inevitably must be put to work in the ordinary language of legal authorities, and do their work in application to cases where such rules often contradict one another. It is the task of resolving these conflicts that affords legal actors a significant degree of *formal*—rather than merely substantive—discretion.

At this point we begin to approach the limits of the set-theoretic model. The model is, I think, a valuable tool for describing *how legal actors do what they do*. It is unhelpful, and makes no claims to be helpful, in describing what legal actors *should* do, or *why* they should do so, or what *motivates* them. For example, the model gives us no guidance as to which court—the District Court or the Second Circuit—

had the better of the *formal* argument in *Louboutin*.²⁵⁴ It does not tell us whether the *Kremen* court should have “toe[d] the line” of the California Supreme Court’s articulated characterization of the law of conversion or instead built a new predicate around the extension of California cases resolving claims of conversion of intangible property.²⁵⁵ On either of these questions, some more substantive theory—indeed, some *prescriptive* theory—of the role of judges in a legal system is required.

Such prescriptive theories are legion; they range from Justice Scalia’s “law of rules” to Cass Sunstein’s “minimalism.”²⁵⁶ But these theories do not have to do with the *substance* of legal rules. Nor have they anything to do with *precision* of legal rules—a dimension of linguistic indeterminacy that lacks this kind of normative freight. Rather, they are normative positions regarding the judge’s role in tending to the *relations among legal sets*. And the observation that such normative positions can themselves constrain legal actors’ behavior—if at all—only as meta-rules that share all the set-theoretic qualities of doctrinal rules suggests that the problem of discretion is one of infinite regress. As a matter of logical structure, law is sets “all the way down.”²⁵⁷ Later-in-time legal actors will be constrained by doctrinal rules formulated in advance precisely to the extent that they share the would-be constrainer’s selection and application of the appropriate meta-rules to evaluate against a particular state of the world, such that both actors would construct the same universe of discourse when faced with the same case. Discretionary and possibly strategic selection of membership or inclusion moves in the construction of a universe of discourse gives a legal actor significant power, not only to determine adjudicative outcomes, but also to shape the development of doctrine in ways that other legal actors at other points in time might disagree with. Conversely, however, each such act of discretion will influence future universes of discourse exactly to the extent that future legal actors allow

²⁵⁴ See *supra* Section IV.C.

²⁵⁵ See *supra* note 96 and accompanying text.

²⁵⁶ See generally Hathaway, *supra* note 238.

²⁵⁷ John Robert Ross, Constraints on Variables in Syntax (Sept. 1967) (unpublished thesis, Massachusetts Institute of Technology) (on file with DSpace@MIT, Massachusetts Institute of Technology) (“‘[W]e live on a crust of earth which is on the back of a giant turtle.’ . . . ‘But what does this second turtle stand on?’ persisted James patiently. To this, the little old lady crowed triumphantly, ‘It’s no use, Mr. James—it’s turtles all the way down.’”).

them to. Constraint, in this view, is no more or less than convergence of the logical moves underlying the behaviors of legal actors over time. In this view, discretion is—as realists claim—an ineliminable property of *any* system of category-based rules numbering more than one, while legal determinacy is—as positivists claim—a matter of social fact—of *practice*—rather than conceptual, logical, or moral necessity.

CONCLUSION

In this Article, I have developed a formal and structural model of legal systems based on the tools of naïve set theory. I have demonstrated how the tools of set theory may be applied to legal authorities and the behavior of legal actors. I have shown how the structure of systems built on set-theoretic logic gives rise to subtle formal distinctions in the set-theoretic relations among which legal actors must choose in the work of advocacy and adjudication. And I have demonstrated how recognition of this deep formal structure of legal systems both clarifies the terms of debate in various areas of legal theory and implies some emergent structural properties of legal systems that substantive legal theory must grapple with.