

# AA-RM WRESTLING: COMPARING ANALOGICAL APPROACHES AND RULE MODELS FOR LEGAL REASONING

Adam Rigoni 

Barrett Honors College, Arizona State University,  
Tempe, Arizona, USA

## ABSTRACT

Legal reasoning is commonly thought of as being based on either rules or analogies. More specifically, there is ongoing debate regarding whether precedential reasoning is best characterized as rule-based or analogical. This article continues that work by comparing recent and representative approaches from each camp, namely, Stevens’s analogical model and the “rule-based” model of Horthy and Rigoni. In the course of the comparison improvements on each approach are suggested and the improved models serve as the basis for the ultimate evaluation. The evaluation demonstrates that the “best” approach depends on the goals one has in theorizing legal reasoning as well as the jurisprudential assumptions one is willing to make.

---

## I. INTRODUCTION

Two recent articles have advanced the debate regarding whether precedential reasoning is best characterized as rule-based or analogical.<sup>1</sup> This article continues that work by comparing recent and representative approaches from each camp, namely, Stevens’s analogical model<sup>2</sup> and the “rule-based” model of Horthy<sup>3</sup> and

1. Katie Atkinson & Trevor Bench-Capon, *Reasoning with Legal Cases: Analogy or Rule Application?*, in PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND THE LAW 10 (2019); Katharina Stevens, *Reasoning by Precedent—Between Rules and Analogies*, 24 LEGAL THEORY 216 (2018).

2. Stevens, *supra* note 1.

3. John F. Horthy, *Rules and Reasons in the Theory of Precedent*, 17 LEGAL THEORY 1 (2011); John F. Horthy & Trevor J.M. Bench-Capon, *A Factor-Based Definition of Precedential Constraint*, 20 A.I. & L. 181 (2012).

Rigoni.<sup>4</sup> That comparison requires clarifying both the language used to describe models of precedential reasoning, which is taken up in Section II, and the standards for evaluating such models, which are taken up in Section III. With those clarifications made, the models can be compared. The two representative approaches are explained in Section IV, and the comparisons are undertaken in Section V. In the course of the comparison improvements on each approach are suggested and the improved models serve as the basis for the ultimate evaluation. The goal is to provide an assessment of each approach in its best form. The hope is that doing so will illustrate the strengths and weaknesses of analogical and rule-based approaches to legal reasoning in general.

## II. THE TERMS OF THE DEBATE: WHAT MAKES A MODEL “RULE-BASED”?

I’ve written previously about rule-based approaches using a definition from Sherwin and Alexander that identified “rule-based” models as involving indefeasible (strict), outcome-determinative rules extracted entirely from individual past cases.<sup>5</sup> That definition is fine for some purposes, but hardly universally accepted. For example, both Stevens<sup>6</sup> and Atkinson and Bench-Capon<sup>7</sup> classify some models using defeasible rules as rule-based. As legal philosophers have begun to appreciate the relevance of work in logic and in AI and law, I think the term “rule-based” threatens to create more confusion than it’s worth. Hunter’s 2003 article on analogy offers a good example of this confusion.<sup>8</sup> He criticizes Brewer’s work from 1996<sup>9</sup> for its assertion that research in AI and law supports the view that analogical reasoning can be represented using rules. He writes,

[Brewer] furthermore incorrectly relies on artificial intelligence research as bolstering his argument that analogy can be performed entirely by rule-based reasoning. . . . He suggests that the work of a number of artificial intelligence researchers, including Ashley and Rissland, shows that analogy is done by rules in artificial intelligence. He says, for example, “Studies of analogy in other fields support the claim that analogy can be represented by a rule-based model. That claim is a fundamental methodological assumption of studies of analogy in the field of artificial intelligence.” The assertion is incorrect,

4. Adam Rigoni, *Common-Law Judicial Reasoning and Analogy*, 20 LEGAL THEORY 133 (2014) [hereinafter Rigoni, *Common-Law*]; Adam Rigoni, *An Improved Factor Based Approach to Precedential Constraint*, 23 A.I. & L. 133 (2015) [hereinafter Rigoni, *An Improved Factor*]; Adam Rigoni, *Representing Dimensions Within the Reason Model of Precedent*, 26 A.I. & L. 1 (2018) [hereinafter Rigoni, *Representing Dimensions*].

5. Rigoni, *Common-Law*, *supra* note 4, at 133, citing LARRY ALEXANDER & EMILY SHERWIN, DEMYSTIFYING LEGAL REASONING (2008), at 75–76.

6. Stevens, *supra* note 1, at §3.

7. Atkinson & Bench-Capon, *supra* note 1.

8. Dan Hunter, *Reason Is Too Large: Analogy and Precedent in Law*, 50 EMORY L.J. 1197 (2001).

9. Scott Brewer, *Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy*, 109 HARV. L. REV. 923 (1996).

as both of the researchers he mentions rely explicitly on case-based reasoning to undertake analogical reasoning in law and at no stage represent their respective domains using rules.<sup>10</sup>

In a sense of “rule-based,” Brewer is correct because any AI model of analogical reasoning at all is going to make use of rules. Holyoak and Thagard’s multiple-constraint model of analogy, which Hunter employs, is itself a set of rules saying that certain mappings are attempted between various aspects of the source and target and of those mappings the most preferred are the ones that satisfy certain constraints.<sup>11</sup> In this sense of “rule-based,” nearly any model of process at all is going to be rule-based.<sup>12</sup> However, there are other senses of “rule-based” in which Hunter’s criticism seems apt. For example, Hunter is correct that none of the AI researchers cited by Brewer represent their domains using rules; if “rule-based” means that the domain is represented using rules, then Hunter has a point. Still, Hunter’s definition does not require that the rules be indefeasible, so it differs significantly from other definitions, such as Alexander and Sherwin’s.<sup>13</sup>

The irony of this example further illustrates the confusion around the term “rule-based.” On the one hand, Brewer does seem to be misunderstanding or mischaracterizing the AI and law work that he cites, which is what was available before 1996 (the year his article was published). At that time case-based reasoning (CBR) was the dominant approach in AI and law, and it represented cases using sets of factors rather than rules, while Brewer’s approach uses cases to generate explicit rules that are then applied.<sup>14</sup> On the other hand, in 1998 Prakken and Sartor showed how CBR representations can be converted into rules, namely, defeasible rules using sets of factors as their antecedents.<sup>15</sup> So, when Hunter’s article was published in 2003, the AI and law literature did support the assertion that you can represent “analogical reasoning” using rules.

10. Hunter, *supra* note 8, at 240

11. Keith J. Holyoak & Paul Thagard, *Analogical Mapping by Constraint Satisfaction*, 13 COGNITIVE SCI. 295 (1989); Keith J. Holyoak, Dedre Gentner & Boicho N. Kokinov, *Introduction: The Place of Analogy in Cognition*, in THE ANALOGICAL MIND: PERSPECTIVES FROM COGNITIVE SCIENCE 1 (Keith J. Holyoak, Dedre Gentner & Boicho N. Kokinov eds., 2001); Keith J. Holyoak, *Analogy*, in THE CAMBRIDGE HANDBOOK OF THINKING AND REASONING 117 (Keith J. Holyoak & Robert G. Morrison eds., 2005).

12. See Rigoni, *Common-Law*, *supra* note 4, at 148 (“One might worry that once we articulate the analogical process, we have reduced it to rules. However, every form of reasoning goes from a set, P, of premises or evidence to a conclusion, C. Hence any form of reasoning could be articulated by a rule such as ‘from P, infer C.’ Yet we are not all rule theorists.”).

13. LARRY ALEXANDER & EMILY SHERWIN, *DEMISTIFYING LEGAL REASONING* (2008), at 75–76.

14. See Atkinson & Bench-Capon, *supra* note 1, at §3, for an excellent summary. Another good introduction with more detail is Trevor J.M. Bench-Capon, *HYPOTHESIS: Introduction to the Virtual Special Issue*, 25 A.I. & L. 205 (2017). The seminal work is KEVIN D. ASHLEY, *MODELING LEGAL ARGUMENT: REASONING WITH CASES AND HYPOTHETICALS* (1990).

15. Henry Prakken & Giovanni Sartor, *Modeling Reasoning with Precedents in a Formal Dialogue Game*, 6 A.I. & L. 231 (1998). See discussion in Atkinson & Bench-Capon, *supra* note 1, at §3.

None of this is meant to imply that a fruitful distinction cannot be drawn between analogical and rule-based theories, once those terms have been explicitly defined. For example, Stevens argues that the distinction between well-functioning theories in each category is this: “The rule-approach’s basic structure is to pick out categorized aspects for the antecedent of a rule from the precedent that then are applied to the present-case. The analogy-approach’s basic structure is the interaction of the two cases in the mind of the reasoner through the mapping.”<sup>16</sup> This is a distinction worth drawing and one that will return in [Section V](#).

However, any attempt at explicit definition in this article is likely to seem artificial given the lack of consensus regarding the terms. Further, I do not wish to add to the confusion swirling around the term “rule-based” by adding yet another definition. For these reasons, I adopt a different strategy in what follows: I select the most promising approach from those that identify as rule-based and the most promising one from those that identify as analogical, and then treat those as representative of the entire classes. The idea is to look at the differences between the best contenders from each class of theories, but I don’t justify my selection of contenders. A full comparison between all the theories that are identified as “rule-based” or “analogical” would certainly be useful, but would go beyond my purpose here, which is to examine the differences between the views. The interested reader is referred to the arguments in favor of each view found in their respective articles.<sup>17</sup>

### III. WHAT’S A MODEL GOOD FOR?

The question of which of the two models is better, simpler, is difficult to answer. Models are created for various purposes and evaluations must take that into account. For example, in the context of legal reasoning machine learning techniques are typically used to generate predictions of decisions but not to explain or describe the underlying reasoning process.<sup>18</sup> Extracting explanatory insight from such models can be extremely difficult. Consider the machine learning program of Aletras et al., which explains its predictions via lists of high-frequency words from opinions where the European Court of Human Rights found a human rights violation.<sup>19</sup> The list for cases finding violations of rights regarding the living conditions of detainees reads: “prison, detainee, visit, well, regard, cpt, access, food, situation, problem, remained, living, support, visited, establishment, standard,

16. Stevens, *supra* note 1, at 248.

17. See Stevens, *supra* note 1; Rigoni, *Common-Law*, *supra* note 4; Rigoni, *An Improved Factor*, *supra* note 4; Horty, *supra* note 3; Horty & Bench-Capon, *supra* note 3.

18. See Atkinson & Bench-Capon, *supra* note 1.

19. Nikolaos Aletras et al., *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, 2 PEERJ COMPUT. SCI. e93 (2016), <https://peerj.com/articles/cs-93>.

admissibility merit, overcrowding, contact, good.”<sup>20</sup> One struggles to see how this helps us understand the reasoning process of the judges in such cases.

Some theories of legal reasoning, such as Hunter’s, are explicitly “only” descriptive. He writes, “this model of legal analogical inference is only descriptive. It only looks to what judges actually do, and does not concern itself with what they should do. The model is therefore in no way prescriptive or normative: it is only a description of analogical reasoning in law.”<sup>21</sup> A more thoroughgoing prescriptive theory is Dworkin’s interpretivism, at least on some readings.<sup>22</sup> However, as I’ve argued elsewhere,<sup>23</sup> no model can be purely descriptive or purely normative. Briefly, on the descriptive side the modeler still must make decisions to idealize or prioritize the data to be explained. On the normative side, in order to be a theory of an actual practice, some aspects of the actual practice must constrain the theory.

Stevens characterizes the goals of her model as follows:

First, I want to ensure that I can plausibly claim to describe the legal practice of reasoning by precedent, and not some other hypothetical form of reasoning that involves past cases. Therefore, I will assume that an adequate account must be able to meet some minimal descriptive requirements. The account should be able to integrate important aspects like following and distinguishing as central features of reasoning by precedent. In addition it should also integrate the central place of precedent-opinions and rationes decidendi in the practice of reasoning by precedent, and it should be sensitive to the ways opinions and rationes are usually formulated. Second, I am not interested in providing an account that describes how judges actually come to their decisions in the real world. Empirical work is better suited for this task. Rather, I will attempt to show that there are both a rule-approach and an analogy-approach of reasoning by precedent that fulfill minimal descriptive requirements and that also fulfill the normative requirement that reasoning according to these approaches can reliably lead to justified decisions.<sup>24</sup>

I’ve explained my purpose similarly, though less clearly, as offering a theory at the descriptive end of the “normative/descriptive spectrum” that aims to explain judicial reasoning in the language of cognitive science/philosophy.<sup>25</sup>

20. *Id.* at tbl. 3. See also Atkinson & Bench-Capon, *supra* note 1, at 9 (pointing out that in legal reasoning, machine learning techniques tend to be significantly less accurate than systems using handcrafted representations). See Trevor Bench-Capon, *The Need for Good Old Fashioned AI and Law*, in INTERNATIONAL TRENDS IN LEGAL INFORMATICS: A Festschrift for ERICH SCHWEIGHOFER 23 (Walter Hotzendorfer, Christof Tsocholl & Franz Kummer eds., 2020).

21. Hunter, *supra* note 8, at 1228.

22. RONALD DWORKIN, *LAW’S EMPIRE* (1986).

23. Rigoni, *Common-Law*, *supra* note 4, at 135–136.

24. Stevens, *supra* note 1, at 218–219. She goes on to argue that the analogy approach offers normative advantages. Stevens, *supra* note 1, at 249–254.

25. Rigoni, *Common-Law*, *supra* note 4, at 135.

The common strain is an attempt to help *explain* the real-world phenomena of judicial reasoning by presenting it as a genuine kind of reasoning—not as an agglomeration of attitudes, or weighing of key words, or sequence of neurons firing, but the sort of process that at least *could* meet Stevens’s normative requirement that decisions be justified.

This overlap in purpose between Stevens and Rigoni provides a basis for the evaluation of the two approaches in Section V. However, even Stevens’s precise description only offers a rough guide. As she says, both rule-based and analogical theories can meet the minimal descriptive requirements, but these are only *minimal* requirements. What would make the model more than minimally descriptively accurate is left open. Likewise, it’s possible that one model bests the other with respect to simplicity or explanatory power. It’s also possible that one model offers better justification for the decisions reached through the reasoning, as Stevens claims the analogical model does.<sup>26</sup> As the next sections show, the interesting differences between the models force us to consider these other, left-open standards of evaluation. A complete evaluation is then necessarily conditional on an understanding of the relevant theoretical virtues or standards for accuracy. I try to make this conditionality explicit throughout the comparison, but it’s important to note up front.

That said, it’s also worthwhile to explain how I understand some of these standards, though I won’t be arguing for my view here. I count it as a reason in favor of a theory if it can be expanded to cover a broader swath of legal reasoning than it was initially designed for. I think that whether a model is computationally tractable is relevant to descriptive accuracy, because the brain is computationally limited, and because it means the model could be empirically tested to some degree. AI and law offers the best examples of such testing, where large sets of cases are converted into usable representations and different models can be tested to see how well they predict the outcomes.<sup>27</sup> One can find a single illustrative case for just about any theory of legal reasoning. Testing the theories against a case base with multiple cases is much less ad hoc, though not without potential problems.<sup>28</sup>

26. Stevens, *supra* note 1, at 249–254.

27. See, e.g., Stefanie Brüninghaus & Kevin D. Ashley, *Finding Factors: Learning to Classify Case Opinions Under Abstract Fact Categories*, in PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 123 (1997); Latifa Al-Abdulkarim, Katie Atkinson & Trevor Bench-Capon, *A Methodology for Designing Systems to Reason with Legal Cases Using Abstract Dialectical Frameworks*, 24 A.I. & L. 1 (2016); Kevin D. Ashley & Stefanie Brüninghaus, *Automatically Classifying Case Texts and Predicting Outcomes*, 17 A.I. & L. 125 (2009); Vincent Aleven & Kevin D. Ashley, *Evaluating a Learning Environment for Case-Based Argumentation Skills*, in THE PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW (1997).

28. For example, the encoding of the cases so they can be input into the model is often done by hand and offers the modeler a chance to rig the game.

#### IV. THE REPRESENTATIVE THEORIES: THE REASONS MODEL AND STEVENS'S ANALOGICAL ACCOUNT

Here I introduce the two representative theories. Representing the rule-based side is the reasons model (RM), which was developed by Horty<sup>29</sup> and continues to be expanded and refined by Horty and Bench-Capon,<sup>30</sup> Rigoni,<sup>31</sup> and Broughton.<sup>32</sup> The analogical side is represented by Stevens's analogical approach (AA), which makes use of the multiple-constraint model of Holyoak and Thagard.<sup>33</sup>

##### A. The RM

A fairly nontechnical introduction the RM is found in Rigoni's 2015 article, which is the basis for the following characterization. Building on Ashley<sup>34</sup> (1991), and more particularly Alevén,<sup>35</sup> the RM divides a case into four components: (1) factors/reasons in favor of the plaintiff, which we will denote "P $n$ " where  $n$  is a number used to differentiate reasons for the plaintiff; (2) reasons in favor of the defendant, which we will denote "D $n$ " where  $n$  is a number used to differentiate reasons for the defendant; (3) an outcome, which we will denote with "OP" when it favors the plaintiff and "OD" when it favors the defendant; (4) a rule (originally only one), which we will denote "Rule  $n$ " where  $n$  is a number used to differentiate different rules. These components are extracted from the opinions of past cases, but the RM does not specify the process of extraction.

The form of the rule is a conditional created using the other three components. The consequent of the conditional is the outcome from the case. The antecedent of the conditional is a subset of the reasons for the winning side. For example, a rule may look like this: {P1, P2, P3}  $\rightarrow$  OP, which says "if these three reasons for the plaintiff obtain, then the outcome favors the plaintiff." A rule may not look like this: {D1, P1}  $\rightarrow$  OP, because not all the reasons in the antecedent are not reasons for the prevailing party.

We will soon discuss what happens when these rules conflict and what role nonmonotonicity plays, but first an example is helpful. Suppose there was an oral agreement between the plaintiff and the defendant that the defendant would provide fifty widgets on July 4 to the plaintiff in

29. Horty, *supra* note 3; John Horty, *Reasoning with Dimensions and Magnitudes*, 27 A.I. & L. 309 (2019).

30. Horty & Bench-Capon, *supra* note 3.

31. Rigoni, *An Improved Factor*, *supra* note 4; Rigoni, *Representing Dimensions*, *supra* note 4.

32. Gabriel L. Broughton, *Vertical Precedents in Formal Models of Precedential Constraint*, 27 A.I. & L. 253 (2019).

33. Stevens, *supra* note 1.

34. Kevin D. Ashley, *Reasoning with Cases and Hypotheticals in Hypo*, 34 INT'L. J. MAN-MACHINE STUD. 753 (1991).

35. Vincent Alevén, *Teaching Case-Based Argumentation Through a Model and Examples* (1997) (Ph.D. dissertation, University of Pittsburgh).



exchange for \$50. The plaintiff pays the defendant and the defendant fails to deliver the widgets. The plaintiff sues for specific enforcement and prevails. In this highly simplified case, we can say that the oral agreement is a reason in the plaintiff's favor. From this case we then get the following rule in which P1 is the presence of an oral agreement:

Rule 1: {P1}  $\rightarrow$  OP

The precedential import of this bare-bones case is simply Rule 1, saying that if there is an oral agreement, then rule for the plaintiff (enforce the contract). If this rule were understood monotonically we would run into obvious problems, as it would require future courts to enforce every oral agreement. Yet, on the RM rules are understood as defaults, that is, rules that can be overridden in exceptional cases.

This naturally raises the question of when a default can be overridden. If Rule 1 can be overridden at any time, it does not seem to constrain future judges and hence cannot be precedential. The heart of the theory lies in its notion of constraint. From each case we extract not only a rule but a weighing of reasons. Given the rule, which incorporates the outcome, we can see that those reasons for the prevailing party in the antecedent were deemed to outweigh all the reasons favoring the losing party. Let " $>$ " denote this relation of outweighing. This weighing of reasons is binding, so future courts may not alter it. A rule may only be overridden if the current case involves a novel set of opposing reasons, i.e., a set of reasons that both opposes the party that wins according to the rule and is not a subset of the set of reasons previously outweighed by the reasons in the rule.

In our simple example, we can see that {P1} outweighs the empty set of reasons for the defendant. That is, {P1}  $>$   $\emptyset$ . Its precedential force is thus very weak (trivial, in fact). Rule 1 must be followed only when P1 obtains and there are no reasons in favor of the defendant. Now let's make the example more complicated. Suppose that widget components greatly increase in price after the oral agreement. It will now cost the defendant \$10,000 to make the widgets. This is a reason (undue hardship) in the defendant's favor, call it "D1." Suppose the outcome of the case remains in favor of the plaintiff. The rule in this example is still Rule 1, but the precedential force is stronger than it was in the first example. This case tells us that {P1}  $>$  {D1}. Given this case, Rule 1 must be followed in future cases where P1 obtains (there is an oral agreement) and either there are no reasons in favor of the defendant or the only reason in favor of the defendant is D1.

Suppose a new case comes along with the same facts, except in the interim between the previous case and the new agreement widgets were declared illegal. This novel reason in favor of the defendant (D2) means the judge in this case is not bound to follow Rule 1. She may distinguish



this case on the basis of this reason. If she does so, then she introduces the following new default rule:

Rule 2: {D1, D2} → OD

Rule 2 trumps Rule 1, which is accommodated in the logic by assigning Rule 2 a higher priority than Rule 1. Note that Rule 1 is not deleted and replaced by Rule 2; it simply does not apply when Rule 2 does. This is how the theory captures the difference between distinguishing a precedent and overruling it. Distinguishing occurs when a rule is trumped, while overruling occurs when the rule is deleted and replaced.

The judge's decision to apply Rule 2 introduces the weighing {D1, D2} > {P1} corresponding to the preference for Rule 2 over Rule 1. Judges in future cases now have to abide by this weighing, as well as the weighing from the older case, namely, {P1} > {D1}. In this way more and more relative weights are established as cases are decided. As more weights are established the number of novel sets of reasons decreases and future judges become more constrained.

One might object that presenting distinguishing in terms of the introduction of a new rule that defeats the existing rule strays from the common understanding lawyers have of distinguishing. The idea is that lawyers think of distinguishing as giving more content to a doctrine and providing a more elaborate statement of the rule from the past case. However, I am not sure this presents a significant problem for the RM. First, it's worth noting that, assuming that holdings are rules, a case that distinguishes a past case has to have a different rule than the past case, because the distinguishing case will have a different result. Simply stating that the past rule does not govern the present case does not produce any result at all—we need (given our assumption) some rule that does apply to give us a result. As lawyers know, more than one rule can get you to the same result and hence they ought to know that the inapplicability of one rule does not guarantee any particular result. If lawyers do think that past cases have rules but also think that distinguishing cases do not introduce new ones, then I'm not sure those thoughts are consistent.

Since the supposed problem cannot stem from the introduction of a new rule, perhaps it stems from the failure of the distinguishing case to modify the rule from the past case. However, the distinguishing case cannot be understood as literally modifying the rule from the past case, for a few reasons. First, this would permit a lower court to modify a rule from a higher court and thereby bind courts that are not below it in a judicial hierarchy. For example, a trial court could distinguish a Supreme Court case modifying the rule from the Supreme Court case, and the modified rule would then bind appellate courts. This, of course, is not what happens. The "elaboration" of the distinguishing court only binds courts beneath it in the hierarchy. The same problem can be raised by considering why it is acceptable

for two courts at the same level of a judicial hierarchy to distinguish a higher court opinion in inconsistent ways. For the sake of completeness, I should point out that not only would the elaboration alter the past case, but it would have to alter all previous cases following the original rule as well. Second, modifying the original rule does not comport with typical citation practice. If the original rule is being consistently modified with added exceptions, then one should be able to cite the original case to support a finding that a case falls into the exception. Instead, lawyers cite the case that provided the relevant “elaboration.” Again, the alleged intuition seems inconsistent with other aspects of legal practice.

However, it may not be fair to think of the objection as claiming that the rule in the past case is literally modified. Rather, the idea may be that elaborating doctrine is created without modifying the rule of the past case. The objection is that this type of doctrine is not conceived in terms of rules trumping the original one. While I agree that most lawyers do not think of doctrine in the technical terms of the prioritized default rules of the RM, they sometimes think of it in terms of general rules or principles with more specific rules providing exceptions and clarifying terminology. Witness the Restatement Third, Restitution and Unjust Enrichment, which begins with “A person who is unjustly enriched at the expense of another is subject to liability in restitution.”<sup>36</sup> Later it states that recovery in restitution to which an innocent claimant would be entitled may be denied because of that claimant’s inequitable conduct in the transaction.<sup>37</sup>

None of this is to imply that doctrine is (or should be) *only* thought of in terms of prioritized rules. Although I will later argue that the RM has an advantage over the AA insofar as it offers some insight into how precedent can generate doctrine, doctrine stems from more than just the aggregation of precedent cases. For example, doctrine also comes from theories constructed to account for those sometimes conflicting cases, as one finds in Restatements. My point is just that common understandings of legal doctrine already use prioritized rules, so using them to represent distinguishing cases should not bother common lawyers as much as the objection alleges. One could modify the RM to allow for rules that attribute holdings to past cases,<sup>38</sup> but these would still be new rules created in the distinguishing case and hence would not avoid the problem.

Speaking of modifications, the RM has been modified to handle other aspects of case-based reasoning, including allowing cases to introduce multiple rules of precedent, in particular, rules based around a factor hierarchy.<sup>39</sup> These are used by the RM to represent reasoning about the

36. Restatement (Third) of Restitution and Unjust Enrichment 1 (AM. L. INST. 2011).

37. *Id.* at 63.

38. See Matthew Carey, Holdings About Holdings: Modeling Contradictions in Judicial Precedent, 21 A.I. & L. 341 (2013).

39. See Rigoni, *An Improved Factor*, *supra* note 4; Rigoni, *Representing Dimensions*, *supra* note 4, at §4.2.

presence or absence of certain factors. The hierarchies allow the RM to represent rules, using lower-level factors, for determining the presence of a higher-level factor.<sup>40</sup> For example, the factor P1 (an oral agreement) from our hypothetical contract case could be treated as a higher-level factor in a hierarchy. The lower-level factors would be the presence of an offer, consideration, and acceptance, respectively. We would have a default rule that takes those factors and infers the presence of an agreement. That is,

Rule 3: {Offer, Consideration, Acceptance} → P1

Notably, it's often easier to think of these lower-level factors as favoring or opposing the relevant higher-order factor rather than favoring one party or the other. So instead of "Offer" favoring the plaintiff or defendant, it favors the presence of an agreement (P1). A lower-level factor like "the offer is for the sale of illegal goods" would be a factor opposed to P1.<sup>41</sup> The process of constructing a weighing between opposing factors is the same as before, except it occurs on each level of the factor hierarchy. This modification will come up again in [Section V](#).

Other modifications include proposals for dealing with dimensions—factors that occur on a kind of scale of support or opposition to a party or claim.<sup>42</sup> Consider the speed of the defendant in an action for negligent driving. Speeds above the speed limit favor the plaintiff, with higher speeds being better for him than lower ones. Speeds at or below the speed limit favor the defendant, with slower speeds being more favorable.<sup>43</sup> Here, the speed limit itself is the "switching point," where speeds at or below it favor one party and speeds above it favor the other side. Additional modifications have been suggested to accommodate tiered court structures.<sup>44</sup> The details of these proposals need not concern us in what follows.

In sum, the RM offers a model of legal reasoning in which ratios (and, arguably, other legal rules from cases) are default rules that can only be

40. Stefanie Brüninghaus & Kevin D. Ashley, *Finding Factors: Learning to Classify Case Opinions Under Abstract Fact Categories*, in PROCEEDINGS OF THE 6TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 123 (1997); Aleven, *supra* note 35; Latifa Al-Abdulkarim, Katie Atkinson & Trevor J.M. Bench-Capon, *From Oral Hearing to Opinion in the U.S. Supreme Court*, in LEGAL KNOWLEDGE AND INFORMATION SYSTEMS—JURIX 2013: THE TWENTY THIRD ANNUAL CONFERENCE 1 (Kevin D. Ashley ed., 2013).

41. See The Restatement (Second) of Contracts 178 (AM. L. INST. 1981).

42. John F. Horty, *Reasoning with Dimensions and Magnitudes*, in PROCEEDINGS OF THE 15TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW 109 (2017); Rigoni, *Representing Dimensions*, *supra* note 4; Edwina L. Rissland & Kevin D. Ashley, *A Note on Dimensions and Factors*, 10 A.I. & L. 65 (2002); Katie Atkinson et al., *Argumentation Schemes for Reasoning About Factors with Dimensions*, in PROCEEDINGS OF 26TH INTERNATIONAL CONFERENCE ON LEGAL KNOWLEDGE AND INFORMATION SYSTEMS (JURIX 2013) 39 (Kevin D. Ashley ed., 2013).

43. This example nicely demonstrates that the magnitudes need not nicely follow a number line. It's plausible that beneath some low speed, the fact of his speed would again oppose the defendant because he would be driving dangerously slowly. I ignore this in the main text for simplicity.

44. Broughton, *supra* note 32.

overridden in the presence of a novel reason/factor. The weighing system for reasons provides the fixed content from past cases that binds future judges. Overriding the rule when novel reasons are present captures the phenomena of distinguishing. The theory can be modified to handle inferences from lower-level (more factual or particular) factors to higher-level (more legal or abstract) ones. It's now time to turn to the analogical approach.

## B. The AA

The representative analogical approach (the AA) is the multiple-constraint model presented in Stevens's recent article.<sup>45</sup> The description that follows comes from that work. Stevens's AA uses aspects, argument schemes, and critical questions. Regarding aspects, Stevens writes,

I am here deliberately choosing the vague term "aspect" of a case, which is meant to approximately denote some property/circumstance of, or relation within the case—something the case can be said or has been said to have. That a case has an aspect can be expressed either through precise categorical terms or through vague allusions, and what the aspect is that such a term or allusion refers to might not be clear to the speaker/writer and/or the hearer/reader. If I describe a situation and note "She had a weird *je ne sais quoi*," then I have successfully denoted an aspect. Any situation, case, or object has uncountable aspects.<sup>46</sup>

She describes arguments schemes and critical questions thus:

[Argument schemes] model the premise-conclusion structure of the kind of argument they represent and combine these with a set of critical questions. The critical questions are meant to pick out the types of objections that are most often successfully leveled against this type of argument. A good argument of the type that the argument scheme picks out must be able to withstand these objections.<sup>47</sup>

Stevens's approach begins with the opinion of the past case. The opinion authoritatively establishes the relevance of certain aspects. Those aspects cannot be treated as irrelevant by the current judge. Further, the ratio is composed of those aspects that are established as most relevant to the decision by the opinion. Much like the RM, the AA does not model the extraction of relevant aspects or ratios.

The AA represents the current case as a set of aspects. That set is then mapped against the relevant aspects of the past case as established in the opinion, with the mappings being evaluated according to the multiple-

45. Stevens, *supra* note 1.

46. *Id.* at 20.

47. *Id.* at 221.

constraint theory of analogy.<sup>48</sup> The ratio from the past case provides a minimal standard of similarity; a successful mapping must map an aspect of the present case to every aspect of the ratio.<sup>49</sup> The following argument scheme and critical questions summarize the model:

#### Argument

Precedent case A, as it is presented in the precedent opinion, and present case B are similar.

Therefore, precedent case A and present case B are legally the same.

Whenever a present case B is legally the same as a precedent case A, A must be followed in B.

Precedent case A and present case B are legally the same.

Therefore, A must be followed in B.

#### Questions

Are A and B similar in a legally relevant way?

Can a successful mapping be made to an aspect of the present case for every aspect of the precedent case that the opinion highlights enough to indicate that it is part of the ratio in the opinion?

Does the surrounding law allow present case and precedent case to be mapped successfully?<sup>50</sup>

Are there no legally relevant differences between A and B?

Thus, it seems that precedents bind when there is a successful mapping between aspects of the current case and the ratio of the past case, and no legally relevant differences exist between A and B.

However, Stevens recognizes that some similarities weaken the argument<sup>51</sup> for deciding the present case the same way as the past. To illustrate, suppose we have a past case where plaintiff wanted to void a contract because he signed it as a minor,<sup>52</sup> and the defendant never even inquired about the plaintiff's age. The defendant objected that the plaintiff looks much older than the age of majority. The judge held that despite the defendant's appearance, the defendant is entitled to void the contract. The case can be represented as follows:

Case 1: Aspects = P and D made a contract; P is a minor;

P looks old; D never inquired about P's age.

Holding = P may void the contract.

48. See *supra* note 11.

49. Stevens, *supra* note 1, at 247.

50. I mostly ignore this global requirement in analyzing the models. Stevens notes that we can put an equivalent requirement on the extraction of rules for the RM, see Stevens, *supra* note 1, at 231–233, so it shouldn't affect the comparisons too much. How this requirement constrains the extraction of ratios is too large of an issue to cover here, but it does come up *infra* in Section V.C.

51. The analogical mapping engine can at best make similarities irrelevant, but not harmful.

52. See The Restatement (Second) of Contracts (AM. L. INST. 1981).

Now suppose in the current case the plaintiff is a minor who wants to void a contract. He argues that the defendant never inquired about his age. Further, this plaintiff looks much younger than the age of majority. The current case can be understood as follows:

Case 2: Aspects = P and D made a contract; P is a minor;  
P does not look old; D never inquired about P's age.

The judge looks to the past case to help her decide the current case. The current case, Case 2, is missing an aspect of Case 1, namely, in Case 2 the plaintiff does not look old. At most we could say that this dissimilarity is unimportant—perhaps the appearance of the plaintiff is just irrelevant. However, what we want to say is that Case 2 is easier than Case 1; it's an even better case for the plaintiff because a reason for the defendant has been removed. This issue was one of the main motivations for factor-based representations of cases, since factors carry a polarity for one party or the other.<sup>53</sup>

Hence, Stevens permits distinguishing only when the legally relevant difference (aspect) opposes the outcome of the past case. The gist of the AA model has now been presented, but further clarification on the role of aspects will be helpful before drawing comparisons and highlighting potential benefits of one model or the other. The next subsection gives this clarification.

### C. Clarifying Aspects

In [Section V](#) we will see that many supposed advantages of the AA stem from its use of aspects. It's then worth spending some time clarifying what aspects do within the theory and how they differ from factors. At first blush, aspects seem different from factors insofar as aspects are properties of the case itself while factors are representations of the case constructed outside the model. This is correct, but it is not as important as it might seem.<sup>54</sup> Aspects, as they are used within the AA, are very similar to factors.

53. KEVIN D. ASHLEY, *MOLDING LEGAL ARGUMENT: REASONING WITH CASES AND HYPOTHETICALS* (1990).

54. There is a somewhat technical point regarding whether aspects are different from factors insofar as factors will not admit a magnitude. A single factor is either present or absent but cannot take on a range of values. For example, there is no factor for "speed of the car" in an automobile accident case. E.g., in a case about an automobile accident you could have a factor for "the car's speed was 45 mph" and another for the "the car's speed was 46 mph," but not a single factor "the car's speed" that takes on different magnitudes of speeds. It may be that aspects can take on magnitudes, though I'm not sure that is how Stevens envisions them. In any event, this does not seem to amount to much of a difference. Contrary to some in AI and law, *see* Horty, *supra* note 29; Latifa Al-Abdulkarim, Katie Atkinson & Trevor J.M. Bench-Capon, *Statement Types in Legal Argument*, in *LEGAL KNOWLEDGE AND INFORMATION SYSTEMS—JURIX 2016: THE TWENTY-NINTH ANNUAL CONFERENCE 3* (2016), I advocate for treating these sorts of things as dimensions that are composed of a set of standard factors with an ordering relation. The ontology is still just factors and an ordering relation; the dimensional aspect just adds a background ordering against which the case occurs. *See* Rigoni, *Representing Dimensions*, *supra* note 4.

To show this, we must go deeper into the details of the AA. What follows is a slight formalization of the mapping process that goes beyond what Stevens describes. I think it is an accurate representation of her view, but it's a reconstruction, not a summary, of what she wrote. She writes that the main difference between the RM and the AA is that "[t]he rule-approach's basic structure is to pick out categorized aspects for the antecedent of a rule from the precedent that then are applied to the present-case. The analogy-approach's basic structure is the interaction of the two cases in the mind of the reasoner through the mapping."<sup>55</sup> Per Stevens, on the RM the content of the current case (its factors) and the content of past cases (their factors and ratios) are independent. On the AA, the content of past cases depends on the current case and the content of the current case depends on the past cases. The opinions are converted into aspects and a ratio in light of how similar they are to the current case. The current case can be characterized differently depending on the similarities to the background case. There is no fixed background or current case; each can shift depending on the other.<sup>56</sup>

We begin with the determination of the content of the current case because the role of aspects is clearest there. As mentioned in the previous section, aspects are all the properties of the case. The AA cannot start with the full set of aspects of the current case, because that would be an unmanageable infinite set and the mapping process would never end. There would always be more available mappings that need evaluation in light of the multiple-constraint model, so the model could never select a best mapping (or mappings).

Rather, I think what Stevens has in mind is that the current case starts as a (still huge) subset of its aspects, a subset composed of every possibly relevant aspect of the case.<sup>57</sup> Mappings take subsets from this set and associate each member (each aspect in the subset) with an aspect from the past case. These mappings are evaluated according to the standards and restrictions discussed in Section IV.B. I will call each of these subsets, i.e., each subset of possibly relevant aspects used by a mapping, a "characterization" of the current case. The choice of a characterization of the current case depends on the quality of the mapping between that characterization and the past case. Hence the characterization of the current case depends on the past case. Similarity with the past case guides the judge in her determination of which of the possibly relevant aspects are actually relevant.

55. Stevens, *supra* note 1, at 248.

56. I ignore that both theories will eventually have to address how one case in a case base can effectively reinterpret another. A case may be read a certain way for years, but if a higher court announces a different interpretation, then that becomes the reading lower courts must adopt. I think both views can accommodate this, but the problem is irrelevant here since it applies equally to both theories.

57. I ignore the problem of how to define "possibly relevant aspects" in a way that avoids an equivalence with all aspects.



Explaining the influence of the current case on the past case requires further reconstruction of the AA. In the AA, as noted in [Section IV.B](#), the characterization of the past case depends on the language of its opinion. The opinion provides authoritative descriptions of aspects and the ratio of the case. However, to enable a two-way interaction between the present and past case, the language of the opinion must be pliable enough to yield a number of different acceptable interpretations.<sup>58</sup> Let an interpretation be a pair composed of a set of relevant aspects and a ratio. A past case can then be represented as a set of interpretations. Now the current case is a set of characterizations and the past case is a set of interpretations. Each characterization is mapped onto each interpretation, and the mappings are ordered according to the multiple-constraint model. The judge uses the characterization-interpretation pair (C-I pair) at the top of the ordering. Using a different characterization can result in a different interpretation appearing in the top C-I pair, which is how the content of the background case depends on the current case. Likewise, as discussed above, using a different past case can result in a different characterization appearing in the top C-I pair, which is how the content of the current case depends on the background case. Of course, it's possible multiple C-I pairs will tie and the judge will have discretion in choosing between them.

We can now note a difference between aspects and factors. Some aspects are not possibly relevant, while others may be possibly but not actually relevant. In contrast, factors are always actually relevant, where relevance means their presence or absence strengthens or weakens a conclusion that either decides the case or forms part of a chain of conclusions that decides the case. However, notice that aspects that are not possibly relevant do not play a role in the AA. So in comparing the models the meaningful difference is that aspects can be either merely possibly relevant or actually relevant, while factors are always actually relevant. [Section V](#) examines whether this difference provides the AA an advantage.

We can also notice several similarities between the two. First, both the RM's representation of the current case in terms of factors and the AA's representation of the case in terms of potentially relevant aspects are exogenous to their respective models. Just as the RM gives no procedure for how the current case is converted into factors, the AA gives no procedure for how the current case is converted into a set of possibly relevant aspects. Each theory requires that the initial representation of the current case be exogenous. This is unsurprising, since no model is going to take "raw data" as an input. Consider the real-world raw data that a judge receives.

58. As I will discuss later, the more pliable the opinion is, the greater advantage the AA has over the RM. If we assume that the language of the opinion by itself fixes the actually relevant aspects, then, given my suggested equivalence between aspects and factors, the two models will look very similar.

She has her own past knowledge, and then receives evidence via testimony, visual cues, etc. She must decide how to evaluate and interpret that evidence in order to produce a representation of that case. There are models of legal reasoning using such evidence,<sup>59</sup> but none that start from a recording of testimony and a set of briefs, for example. They all start with the data being converted into a usable form. Further, though the reasoning at this initial stage will likely involve analogy, it will also involve application of other forms of reasoning, such as Bayesian reasoning in assigning credence to testimony. Most importantly, this reasoning is everyday reasoning, not distinctively legal reasoning. There simply are no legal rules or precedents for how much credence a judge should give some bit of testimony. This is illustrated vividly in the American system in the difference between determinations made by a jury, which are obviously not bound by precedent, and those made by a judge.<sup>60</sup>

Second, the actually relevant aspects in the AA are equivalent to factors. Both refer to the properties of the case that matter for its resolution. Two possible objections can be raised against this equivalence. The first is that aspects can represent lower-order (more purely factual) properties, such as the existence of written agreement, while factors can only represent complex (more legal) properties, such as the existence of a binding contract. The thought is that aspects represent the properties used to infer the presence of a factor. This objection fails because there is no restriction on how complex an aspect can be nor any limit on how factual a factor can be. The use of factor hierarchies allows the representation of lower-order factors that are then used to infer the presence of a higher-order factor. To put the point another way, you cannot get any closer to raw data by using relevant aspects than you can with factors.

The second objection is that factors have a polarity while relevant aspects do not. The motive for this objection is again the idea that aspects are closer to raw data because raw data does not come with a polarity attached. Admittedly, the polarization of factors is exogenous to the RM. However, as discussed in [Section IV.B](#), Stevens realizes that only aspects that favor an alternate outcome can permit distinguishing. This means that at some point in the AA aspects must be assigned a polarity and this appears to happen exogenously, just as it does in the RM. Thus, relevant aspects do come with a polarity in the AA and how this polarity was determined is no more explained within the AA than it is in the RM.

59. Vern R. Walker, *A Default-Logic Paradigm for Legal Fact-Finding*, 47 *JURIMETRICS J.* 193 (2007); Floris Bex, Trevor J.M. Bench-Capon & Bart Verheij, *What Makes a Story Plausible? The Need for Precedents*, in *JURIX 2011: THE TWENTY-FOURTH ANNUAL CONFERENCE* 23 (2011); Floris Bex, *An Integrated Theory of Causal Stories and Evidential Arguments*, in *PROCEEDINGS OF THE 15TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW* (2015).

60. See Rigoni, *Representing Dimensions*, *supra* note 4, at §4.2. More generally, see Al-Abdulkarim, Atkinson, and Bench-Capon, *supra* note 54.

To recap: relevant aspects are just factors. This should make us suspicious of claims that one or the other model is superior because of its use of factors or aspects. The only point where the models differ is that the AA starts with the current case as set of potentially relevant aspects (potential factors) while the RM starts with the current case as a set of factors (relevant aspects). That is, we can think of the AA as starting with cases consisting of a set of sets of factors, which is a set of possible interpretations, while the RM starts with a fixed interpretation. This difference allows the AA to model the determination of which potentially relevant aspects count as factors, but nothing more.

## V. THE MODELS COMPARED

In this section I examine the differences between the RM and the AA and consider the alleged advantages of the AA. As in the previous section, many *prima facie* advantages dissolve once we see how the other account can be modified to do the same thing. I begin with the alleged advantages of the AA as detailed by Atkinson and Bench-Capon. I show that all these benefits are either illusory or slight. I then turn to Stevens's own discussion of the advantages of her view. Here I argue that there are some genuine benefits, but that they trade off with benefits for the RM as well. Finally, I explain the alleged advantages of the RM in explaining legal doctrine and the entrenchment of legal rules.

### A. Atkinson and Bench-Capon on the AA's Advantages

To begin, consider the view of Atkinson and Bench-Capon on when analogy may be necessary. Atkinson and Bench-Capon write that the AA is needed over the RM in three situations, but I think the key idea can be reduced to this claim: "It therefore appears that providing arguments for cases where the rules run out . . . must remain the province of human lawyers."<sup>61</sup>

Per Atkinson and Bench-Capon, in cases where "legal rules run out" the judge must have recourse to information outside of the case base.<sup>62</sup> They point out that modeling the determinations of whether a lowest-level aspect (or factor) is present will require going far past any ontology of factors (or aspects) seen in current models. Taking an example from Stevens, they consider how to model a judge addressing the question of whether a kindergarten teacher has a sufficiently close relationship to a student to ground a claim of emotional distress, given a past case holding that mothers have a sufficiently close relationship to children to ground a claim of emotional distress.<sup>63</sup> The model would need the background knowledge about the nature of the relationships between mothers and children and kindergarten

61. Atkinson & Bench-Capon, *supra* note 1, at §8.

62. *Id.* at §8.

63. *Id.*

teachers and students, among other information, that the judge brings to bear on this comparison.<sup>64</sup> Atkinson and Bench-Capon note that what is needed is likely a general-purpose ontology that would cover all the basic concepts and “rules of thumb,” i.e., a representation of common sense.<sup>65</sup>

Initially the claim that the AA is needed here seems a bit puzzling. We just saw that “rules of thumb,” i.e., defeasible rules, lie at the heart of the required ontology. The nonmonotonic techniques that underlie the RM were developed to represent just such rules.<sup>66</sup> However, notice that for Atkinson and Bench-Capon the alternative to the RM is that the process remains with computational modeling in “the province of human lawyers.” Stevens’s AA is built around the system of Holyoak and Thagard, which is computational.<sup>67</sup> Atkinson and Bench-Capon are not interested here in comparing the RM to this computational alternative. Rather, they are interested in pointing out where the computational models need to be supplemented. One might say they are interested in the limits of modeling itself in legal reasoning. This is a different matter from the dispute between the AA and the RM, which are both theoretically computational models.

However, Atkinson and Bench-Capon’s discussion does raise a potential advantage for the AA, in that the AA can help in cases where precedent is not binding but persuasive. That is, when no ratio or ratios are available to cover the current case. To illustrate, consider two past cases and a current one such that

Past Case 1: Factors/Aspects = {P1, P2, P4, D1}; Ratio = {P1, P2} → OP

Past Case 2: Factors/Aspects = {P2, P3, D2, D3, D4}; Ratio = {D2, D3, D4}  
→ OD

Current Case: Factors/Aspects = {P2, P3, P4, D1, D2, D3}; Ratio = ??

Let the factors be defined as follows:

P1: defendant, a business, refused to clear its sidewalk after a previous complaint

64. *Id.*

65. *Id.*

66. GERHARD BREWKA, NONMONOTONIC REASONING: LOGICAL FOUNDATIONS OF COMMONSENSE (1991); Nicholas Asher & Michael Morreau, *Commonsense Entailment: A Modal Theory of Nonmonotonic Reasoning*, in PROCEEDINGS OF THE 12TH IJCAI 387 (Morgan Kaufmann ed., 1991); G. Aldo Antonelli, *Non-Monotonic Logic*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Winter ed. 2012), <http://plato.stanford.edu/archives/win2012/entries/logic-nonmonotonic/>.

67. Keith J. Holyoak & Paul R. Thagard, *A Computational Model of Analogical Problem Solving*, in SIMILARITY AND ANALOGICAL REASONING 242 (Stella Vosniadou & Andrew Ortony eds., 2009).

P2: plaintiff slipped and fell on ice on defendant's sidewalk

P3: other businesses near defendant regularly clear their sidewalks

P4: the ice covered the entire width of the sidewalk

D1: plaintiff was inebriated when he fell

D2: defendant had placed an orange traffic cone in front of the ice on the sidewalk

D3: plaintiff was walking on the sidewalk after business hours

D4: defendant had attempted to clear the ice with salt on the day of the accident

Here neither ratio applies to the present case because it lacks P1 and D4. Thus precedent as conceived of on both the RM and AA does not compel a ruling one way or the other. Still, even if precedent does not require a certain decision, we might think a judge legally ought to decide the case in the same way as the most similar case. Or, we may think that judges do in fact exercise their discretion in this way. If you accept that similarity does or should decide these cases, then it would be a benefit for a theory of legal reasoning to capture this. To be clear, I am not asserting that similarity does or should do this. I am asserting that it's reasonable to think it does, and hence we ought to see if either theory can better accommodate this view.

The question is, can the AA capture this in a way that the RM cannot? The answer is complicated. In a sense, both theories could be modified to include a provision that when no ratio is available, cases should be decided in terms of the number of similarities. Past Case 1 has three factors (P2, P4, D1) in common with the current case, while Past Case 2 has four (P2, P3, D2, D3). Thus, the case should be decided for the defendant. However, this treats all relevant similarities as equally important, which seems wrong. If the size of the ice patch (P4) is really important, then perhaps Case 1 is the most analogous. The AA seems to have an advantage here, because it provides a method for determining which of the two cases is more similar (which mapping is the best) using the constraints (surface, structural, etc.) of the analogical model.

Yet, this ignores that in the AA the past case's determination of which factors are important, as expressed in the ratio of the past case, is part of what binds the current judge.<sup>68</sup> An appropriate modification of the AA to deal with these cases must, for example, treat P1 as important because the judge in Case 1 did, even if the constraints in the mapping engine would

68. Stevens, *supra* note 1, at 242–247.

mark it as unimportant. Extending the requirements on the mappings to accommodate this would not be difficult. Doing so would just add another constraint to the existing ones. The constraint would be something like a stipulation that when no mapping establishes similarities with all aspects of the ratio, mappings that establish some of those similarities are to be favored over those with none. This creates a three-tier hierarchy of mappings where you have those that establish similarity with the whole ratio at the top (strictly most preferred), those that establish some similarities with the ratio in the middle, and those that establish one at the bottom (strictly least preferred). Between mappings within any of these categories the constraints work as usual.<sup>69</sup>

While selecting the most similar past case seems like a domain naturally suited to the AA, the RM can be modified to address this as well. The modified RM may even be superior to that of the AA, because it grounds determinations of similarity in the case base. The modification is to use the weightings of factors that are established in the case base to provide a relative weight for each shared factor (similarity). The RM explicitly provides for comparison in weight between factors of different polarities. There are also implicit comparisons that can be drawn between factors of the same polarity, e.g., if  $R1 > D2$  and  $D2 > R2$ , then  $R1 > R2$ . Magnitudes introduce further weightings between factors of the same polarity.<sup>70</sup> Finally, the RM could adopt an analog of the AA's approach to ratios and stipulate that each factor in a ratio outweighs any one of the other factors in the case.<sup>71</sup> The modified RM would use all these relative weights to give a measure of importance for similarities.<sup>72</sup> The modified RM's standard for decisions in these cases would be this: follow the case with the highest weighted similarities for each side. Returning to our most recent example, suppose that  $P4 > P3$  (a large ice patch is a better reason for the plaintiff than the fact that other businesses clear their sidewalks regularly), and  $D1 > \{D2, D3\}$  (a drunk plaintiff taking a fall is better for the defendant than having traffic cones by the ice and having the fall occur

69. There are probably several technical difficulties this modification would have to surmount. Since I am trying to compare the theories in their best modified forms, I err on the side of optimism, assuming that the details of a modification can be worked out whenever it raises no obvious concerns.

70. See Rigoni, *Representing Dimensions*, *supra* note 4, at 4–9.

71. This strikes me as a good addition to for the RM independent of any comparison with the AA. It would be odd if a court held back the strongest reasons when providing a ratio. The potential context sensitivity of the weight of reasons is a potential problem: the ratio might be the strongest reason relative to the opposing ones present in the case, but not the strongest simpliciter.

72. Another approach would be to tie similarity to the value promoted by the relevant factor so that factors promoting the same value are treated as similar. See Trevor J.M. Bench-Capon & Giovanni Sartor, *A Model of Legal Reasoning with Cases Incorporating Theories and Values*, 150 A.I. 97 (2003). For further discussion of the role of values in weightings, see Alison Chorley & Trevor J.M. Bench-Capon, *An Empirical Investigation of Reasoning with Legal Cases Through Theory Construction and Application*, 13 A.I. & L. 323 (2005).

after hours). This means that the pro-plaintiff factors shared by Past Case 1 and the current case (P2 and P4) are more important than those shared by the current case and Past Case 2 (P2 and P3). In terms of their similarities, Past Case 1 is a stronger case for the plaintiff than Past Case 2. Also, the pro-defendant similarity shared by the current case and Past Case 1 (D1) is more important than both the similarities between the current case and Past Case 2 (D2, D3). In terms of the shared factors (similarities), Past Case 1 represents a stronger case for the defendant than Past Case 2. Informally, we could say that Past Case 2 is the best the judge can do as far as a past case that deals with the strongest reasons for each side.<sup>73</sup>

The modified RM determines similarity in a distinctively legal way that respects the decisions made not just in a past case, but throughout the case base. It explains some of the global coherence that is a desideratum of legal decisions,<sup>74</sup> even when the decision is not strictly compelled by precedent. For that reason, it seems like a better way of deciding cases according to similarity. However, the AA theorist can reply that this advantage only covers a narrow range of cases. Unless the case base is unrealistically extensive, it will not establish a complete ordering of weights for even a single factor. Hence there will still be similarities that the judge must weigh herself. Here the RM adds nothing, while the multiple-constraint model of the AA can operate to weigh similarities. The advantage of the AA model is that some of its constraints are derived from studies of human psychology, and thus they can operate when legal materials run out.<sup>75</sup>

The value of this benefit for the AA depends on how often judges must decide a case using a case base that provides no applicable precedent, but still provides more than one relevant case. Notice that if we suppose our case base is just Past Case 1 instead of Past Case 2, then relative similarity does not matter—the only mappings that are possible all map to Past Case 1. Additionally, the mapping constraints of the AA won't tell us when the prior case is too dissimilar to follow, nor will they help if the past case shares only a few (or no) similarities to the present case, i.e., a case of first impression cannot be resolved via similarity. Plus, the AA does not guarantee that there will be a unique most similar case. It's possible for mappings to tie according to the constraints of the model, which means the judge must select a mapping according to some criteria external to the model. If ties are common, then this supposed advantage amounts to little in practice.

73. As noted regarding the AA in note 69, *supra*, there are a number of potential problems this modification faces. Again, I am being intentionally optimistic.

74. The most famous proponent of global coherence in judicial decisions is Dworkin. See Ronald Dworkin, *Hard Cases, in TAKING RIGHTS SERIOUSLY* 81 (rev. ed. 1978). The most relevant proponent for my purpose is Stevens. See Stevens, *supra* note 1, at 232–234.

75. There is a possible hybrid view that first uses the RM and the modification for weighing similarities and then applies the AA when that fails to resolve the case. I ignore that here since my focus is on comparing the two theories.



Finally, and perhaps more importantly, the AA's advantage rests on the assumption that when legal rules run out cases should be or are decided according to similarity. I think the strongest objection available to the RM theorists is to challenge this assumption. Why not think that when legal materials run out judges decide, or should decide, the case using moral reasoning or policy preferences or some other reasoning beyond similarity? I can't hope to refute or sustain this challenge here, but the debate should center on this issue.

#### B. Stevens on the AA's Advantages

Regarding the differences between the RM and the AA, Stevens writes, “[t]he rule-approach’s basic structure is to pick out categorized aspects for the antecedent of a rule from the precedent that then are applied to the present-case. The analogy-approach’s basic structure is the interaction of the two cases in the mind of the reasoner through the mapping.”<sup>76</sup> According to her, this difference provides the AA with an advantage because it can “help judges to make sure that they pay attention to ways of understanding precedent and present-case that are different from their first, intuitive impressions of the cases.”<sup>77</sup>

As explained in [Section IV.C](#), interpretations of the past case interact with characterizations of the current case to produce the C-I pair that provides the final content of the current case and the past one. I consider separately the potential advantages of each half of this interaction—determining a characterization of the current case and selecting an interpretation of the past case.

The AA offers an account of how the judge determines which of the potentially relevant aspects of the current case are actually relevant, which is an advantage over the RM. How great of an advantage depends in part on the details of how the model understands “potentially relevant.” Remember, the AA does not explain how the initial set of all aspects of the current case is culled into a set of potentially relevant aspects that are then further reduced to the actually relevant ones. In the limit, the set of potentially relevant aspects would just be the set of actually relevant aspects (factors), which would mean the AA starts with factors just like the RM. If the set of potentially relevant aspects is vast, then the AA explains a difficult problem: how to shift through all those properties to select the actually relevant ones.

Unfortunately for the AA, it can only partially explain how the current case gets characterized. So far, the explanation of this part of the AA has largely ignored that the judge must also be aware of relevant dissimilarities between the current and past case, as noted *supra* at IV.C. These dissimilarities are relevant aspects of the current case that cannot be mapped to the

76. Stevens, *supra* note 1, at 248.

77. *Id.*

past case. According to the process used to select a characterization, these dissimilarities should not be part of the characterization of the case, because they make the mapping worse, not better. As far as I can reconstruct, the AA does not explain how potential aspects become dissimilarities. Nevertheless, making the judge select a characterization does force her to consider some of the different ways the current case could be characterized. Stevens's point that forcing the judge to think beyond her initial impression can improve her decision making is well taken.

However, Stevens does not consider that the RM can also force judges to consider different ways the current case could be characterized. The RM can do this through factor hierarchies, which provide rules (sub-rules) for deriving higher-level factors from lower-level ones. If the initial set of factors for the current case contains at least some lower-level factors with applicable sub-rules, then we can derive a limited number of characterizations by applying (or not applying) the sub-rules, which come from past cases or other legal sources.<sup>78</sup> Those sub-rules that apply and cannot be distinguished must be applied in all characterizations, but sub-rules, just like ratios in the RM, can be distinguished. Hence the judge can decide whether to apply the distinguishable rules, with different decisions yielding different characterizations. This means that the RM can require the judge to consider different characterizations of the current case.

In fact, the RM may be better than the AA in this regard. In the AA, the judge evaluates how well each characterization maps to interpretations to determine the best one. The evaluation is holistic in that it compares *whole* characterizations, i.e., *complete* representations of the current case. In the RM, the judge builds a characterization piece by piece, working up from the lower-level factors to various higher-level factors. Hence, she must examine the reasons for and against each inference from a lower-level to a higher-level factor. By the end of the construction process she knows how each factor in the final characterization ended up there. She is forced not only to reconsider her initial intuition regarding the right characterization of the case, but also to deconstruct that intuitive characterization and consider all the determinations that are implicitly made by adopting it.

The advantages of the two approaches with respect to characterizing the current case are difficult to evaluate without more information about the data the theories attempt to model. I do, however, think it is fair to say there is no clear winner on this front. I now turn to the alleged advantage of the AA in interpreting past cases. We've seen that whatever benefits the AA might derive from modeling the selection of a characterization can

78. Of course, unless the facts of the current case are very common in past cases, the judge will also reason her way to higher-order factors using rules that she creates. The point isn't that the RM obviates the need for the judge to make decisions that are not determined by past case. Rather, the point is that the presence of some rules from the past case forces her to consider the reasons for and against applying them, which entails considering multiple characterizations.

probably be imitated by the RM in modeling the construction of a characterization. It is tempting to do something similar and modify the RM to model the construction of the past case. Unfortunately for the RM, this strategy will not work because there is nothing in the RM that can serve to guide the construction of an interpretation. When characterizing the current case, the RM provides rules from past cases that the current judge applies. But in the RM cases provide no rules until they are decided, so the current case offers no rules at all, leaving nothing for the judge to use in constructing an interpretation of the past case.<sup>79</sup>

Therefore, the AA has a *prima facie* advantage in that it can explain how the judge selects (from the set of acceptable interpretations) an interpretation of the past case, and this advantage cannot be replicated by modifying the RM. The value of this explanation depends on how much interpretative work is being done in the untheorized process of extracting the set of acceptable interpretations from an opinion. If this process usually yields a small set of acceptable interpretations, or a set of extremely similar interpretations, then the AA explains only a small part of the interpretative process, and its advantage seems slight. If the untheorized process typically yields a robust set of varied and conflicting interpretations, then the AA explains a significant portion of legal reasoning that the RM ignores.

Beyond the theoretical virtue of explaining a larger portion of legal reasoning than the RM, Stevens argues that the AA has an additional benefit because it forces judges to consider multiple interpretations, which makes them consider more potential reasons relevant to their decision. This is beneficial because it helps “fulfill the normative requirement that reasoning according to these approaches can reliably lead to justified decisions.”<sup>80</sup> Again, the extent of this benefit depends on how many and how different the available interpretations are. If you think there is a great variety of acceptable interpretations for an opinion, then the benefit may be significant. If you think there are only a few permissible interpretations or that they are all very similar, then the benefit is limited.

Even assuming large numbers of varied interpretations, the AA’s approach is not without costs. From the jurisprudential perspective, one could argue that cases in common law systems are best understood as having a fixed ratio. We see evidence in favor of independent ratios in Restatements, hornbooks, and the like, which compile cases and extract rules/principles/ratios from them without having any particular current case in mind. Still, the AA’s proponents can reply that Restatements and so on are not law, however influential they may be. According to this reply, they are merely aids that offer judges and attorneys a rough starting

79. Although one might see some recent work using the ASPIC+ program as an attempt to extend the RM in this way. See Henry Prakken et al., *A Formalization of Argumentation Schemes for Legal Case-Based Reasoning in ASPIC+*, 25 J. LOGIC & COMPUTATION 1141 (2015).

80. Citing Stevens, *supra* note 1, at 244.

point for the actual process of legal reasoning. Yet, even though Restatements and hornbooks are not law, they do seem to employ legal reasoning. Restatements certainly have their critics, but few seem to criticize them on the basis that the reasoning they use is of the wrong type.

A different jurisprudential objection is that the AA employs a problematic concept of law. There seem to be two options for what the law (the case law) is within the AA. One is that the law is the collection of all the various and potentially conflicting acceptable interpretations of all the past cases.<sup>81</sup> I should stress that this by itself does not seem problematic—anyone who acknowledges judicial discretion (as the RM does) must admit that the law is incomplete and indeterminate in places. The difficulty for the AA on this view is that it must explain why the choice of an interpretation from the set of acceptable ones is not discretionary, but rather the judge is obligated to select and apply the interpretation that is most analogous to the current case. It seems like a judge would still be applying the law regardless of which acceptable interpretation she selected, even if more analogous ones were available.<sup>82</sup>

Another option is that the law is the particular interpretation (or interpretations in the AA as modified to reason about portions of cases, see the end of Section IV.C) that is (or are) in fact most similar to the current case. Now the judge who fails to apply the most analogous interpretation is failing to apply the law. On this view the law has a strong particularist flavor. For example, the law governing a vehicular tort claim could depend on whether the vehicle was a car or a motorcycle. This isn't the uncontroversial claim that the law could contain different provisions for cars and motorcycles. The claim is that the whole law itself could change if the facts of the current case were different. Put another way, the law can differ for cases decided at the same time in the same legal system. On this conception the whole law is relative to a case considered as current.<sup>83</sup>

How troublesome these concepts of law are, of course, depends on one's own theory of law. The idea that the law depends on the particular facts of the current case fits nicely with realist conceptions of law, which focus on the law as what judges do. Since we know from a large psychological

81. On this view the law may be more or less indeterminate, but I don't see that as worrisome since many popular theories of law countenance some degree of indeterminacy. See, e.g., Jules L. Coleman, *Negative and Positive Positivism*, 11 J. LEGAL STUD. 139 (1982); Lon L. Fuller, *American Legal Realism*, 82 U. PA. L. REV. 429 (1934).

82. It's tempting to think the AA illustrates a process of clarifying past opinions, but it only does so locally, meaning that in the AA the ambiguous past case is given a fixed interpretation for the purpose of resolving the present case, but that interpretation is not fixed in future cases. Hence the other acceptable interpretations remain viable for judges in other cases. Stevens's requirement that the surrounding law must permit the mapping can do some work here—it may entail that interpretations of a past case in other past cases further restrict (or exhaust) the set of acceptable interpretations. However, how this requirement works is not thoroughly explained since it is not the focus of the AA.

83. In the same vein that modal logicians talk of worlds being possible given a certain world considered as actual.

literature that human (including judicial) decisions are heavily dependent on normatively irrelevant features of context,<sup>84</sup> a legal realist should expect the law to change depending on the facts of the current case. Other theorists, like positivists, might balk at the notion that you cannot say what the law is without considering a particular set of facts. Relatedly, interpretivists may be happy with the idea that a past case lacks determinate legal content until it is compared with a current fact situation, but others, who stress the importance of citizens understanding what law requires,<sup>85</sup> will be less happy with that.

Whatever our assessment from the jurisprudential perspective, using multiple interpretations of past cases is a detriment from a computational perspective. The representation of the case base becomes exponentially more complicated since each case is a set of possible interpretations. However, even narrow ranges of interpretation threaten to make the process intractable. If each case has four interpretations and we have ten cases, there are over a million potential representations of the case base. Thus, a computational perspective favors the RM.

### C. The Alleged Advantages of the RM

Some potential advantages for the RM have already come up in the course of responding to the alleged advantages of the AA. Here I reiterate some of the benefits that motivated the RM in the first place and that seem unique to it. The RM explains how legal rules can become entrenched<sup>86</sup> and how rules may form hierarchies.<sup>87</sup> As far as I can tell, the AA cannot be modified to replicate these advantages.

On the RM, legal rules become entrenched as the same rule defeats more and more opposing rules (or larger and larger sets of factors, since sets of factors with the same polarity are rules in the RM).<sup>88</sup> Cases that “merely” follow the rule from a past case may and often will involve deciding that that rule defeats a novel opposing rule. Such cases constrain future courts by further entrenching the old rule. Although the AA incorporates legal rules (ratios), it is hard to see how it could explain entrenchment. To begin, it could at best give an account of which ratios are entrenched

84. For the human portion, see Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974); Craig R. Fox & Amos Tversky, *Ambiguity Aversion and Comparative Ignorance*, 110 Q. J. ECON. 585 (1995); Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, 39 AM. PSYCH. 341 (1984); Daniel Kahneman & Amos Tversky, *On the Reality of Cognitive Illusions*, 103 PSYCH. REV. 582 (1996). For results specific to legal reasoning, see Mark Kelman, Yuval Rottenstreich & Amos Tversky, *Context-Dependence in Legal Decision Making*, 25 J. LEGAL STUD. 287 (1996); Pedro Bordalo, Nicola Gennaioli & Andrei Shleifer, *Sabience Theory of Judicial Decisions*, 44 J. LEGAL STUD. S7 (2015).

85. Barbara B. Levenbrook, *The Law of the Street*, in NEW ESSAYS ON THE NATURE OF LEGAL REASONING (James Penner & Mark McBride eds., forthcoming).

86. Grant Lamond, *Do Precedents Create Rules?*, 11 LEGAL THEORY 1 (2005); John F. Harty, *Rules and Reasons in the Theory of Precedent*, 17 LEGAL THEORY 1 (2011).

87. Rigoni, *An Improved Factor*, *supra* note 4.

88. See Harty, *supra* note 3; Harty & Bench-Capon, *supra* note 3.

from the perspective of a particular case (considering a particular case as current) since the ratios can change if we consider a different case as current. More importantly, the AA cannot represent the degree of entrenchment. It could count how often a ratio occurs in a case base, but this by itself cannot tell us how entrenched that ratio is. For all we know that ratio could occur frequently because cases with the exact same relevant facts keep happening. That is, the rule may only defeat one opposing rule, but cases triggering just those two rules are very common.

The AA is a model on which precedential constraint is purely an interaction between one past case (or one decision within a past case) and one current case (or one decision in the current case). The RM explains how decisions from multiple past cases (or multiple decisions within the same case) interact to produce precedential constraint. Another way of putting this is that the RM offers some explanation of how legal doctrines, which are the product of multiple cases, can be created and function. The AA recognizes the importance of doctrine in its requirement that “a judge cannot claim that a mapping is successful or unsuccessful if there are parts of the surrounding law that forbid treating the respective aspects in precedent and present-case as the same or as different.”<sup>89</sup> However, it cannot offer much further explanation of this requirement, while the RM does.

As we’ve seen repeatedly, the extent of this benefit depends on several contingencies. If the case base is composed of independent cases with disconnected ratios, then there won’t be much doctrine for the RM to explain. In addition, the RM only offers a partial explanation of legal doctrine. For example, doctrines can develop as courts identify trends from other decisions, but these trends are by definition extrapolations beyond the decisions themselves, which means the RM cannot explain them. There is also a jurisprudential question of whether precedential constraint should be limited to interactions between individual cases, while doctrine is treated as exerting merely persuasive influence. Nonetheless, the RM has a *prima facie* case for a significant advantage here, though these considerations threaten to mitigate it.

## VI. CONCLUSION

The comparison of the models produces no decisive winner, but some conclusions can be drawn. First, though I’ve challenged some of the claims of Atkinson and Bench-Capon,<sup>90</sup> my comparison supports their conclusion that the AA is unlikely to be formalized in a computationally tractable way. Second, the RM does a better job of explaining the influence of doctrine than the AA. Third, the comparison has demonstrated that both the RM and the AA, in different ways, can force judges to reflect on different

89. Stevens, *supra* note 1, at 244.

90. See *supra* Section V.A.

ways of thinking about the current case, which weakens (but does not entirely overcome) Stevens's claimed advantage for the AA. Finally, the comparison brought to light the different jurisprudential assumptions of the models. The RM assumes that past cases have a fixed interpretation regardless of the current case, while the AA assumes that past cases can have different interpretations depending on the current case.<sup>91</sup> I can't hope to adequately defend or critique either of those views here; the important conclusion is that the plausibility of each view affects the models.

The comparison also makes clear that the differences between the two models largely lies in the fine details, which suggests taking a more detailed look at the data being modeled. Some judges may decide cases by finding and following the most similar case in line with the AA (call them "AA-type"), while other judges may try to apply the applicable rules from a range of relevant cases in line with the RM (call them "RM-type"). Some judges may be committed to treating past cases as having a single constant holding in line with the RM (also members of the RM-type), while others may be happy to change their reading of a past case when they consider a new case in line with the AA (also members of the AA-type). It's not obvious that any of either the RM-type or AA-type judges are doing anything wrong, so idealizing the data, as all models must, may not eliminate either type. Obviously one model would do better for each type of judge, but neither model is going to seem better when looking at the data in aggregate. Further, since the two models are so similar, in many cases the different types of judges may be indistinguishable. Similarly, certain styles of reasoning may dominate certain areas of law or certain levels of a judicial hierarchy, but these distinctions will be erased by combining the data and could be hard to spot even when examining each case individually. Thus, I may have begun by asking the wrong question: rather than asking which theory best explains common law judicial reasoning, I should have asked which parts of common law judicial reasoning are best explained by which theory.

While that is a live possibility worth exploring, it's premature to abandon hope for a unified theory of common law judicial reasoning. Neither the AA nor the RM has been tested on an extensive and varied set of common law decisions, so the actual data may well distinguish them. Further, any reasonable suggestions for how to divide the data, i.e., on where the theories come apart *in the data*, must be informed by an understanding of how and where the two models come apart *in theory*, which this article provides.

91. Subtle issues lie under this description. The AA is not just committed to the idea that legal cases can come to be understood differently over time, which obviously occurs. It is committed to allowing legal cases to be understood differently every time they are applied. The issue for the RM is that sometimes interpretations change; the issue for the AA is that they do not seem to change enough.