A Process Model of Legal Argument with Hypotheticals

KEVIN ASHLEY[a], COLLIN LYNCH[b], NIELS PINKWART[c], VINCENT ALEVEN[d]
[a]*University of Pittsburgh School of Law, Learning Research and Development Center*
[a,b]*Graduate Program in Intelligent Systems, Pittsburgh, Pennsylvania, USA*
[c]*Clausthal University of Technology, Department of Informatics, Germany*
[d]*Carnegie Mellon University, HCI Institute, Pittsburgh PA, USA*

Abstract. This paper presents a process model of arguing with hypotheticals and uses it to explain examples of oral arguments before the U.S. Supreme Court that are like those employed in Socratic law teaching. The process model has been partially implemented in the LARGO (Legal ARgument Graph Observer) intelligent tutoring system. The program supports students in diagramming oral argument examples; its feedback on students' diagrammatic reconstructions of the examples enforces the expectations of the process model. The paper presents empirical evidence that features of the argument diagrams made with LARGO are correlated with independent measures of argumentation ability. The examples and empirical results support the model's explanatory and diagnostic utility.

Keywords. Legal argument; argument diagrams; hypothetical reasoning

1. Introduction

Hypothetical reasoning is a skill of "thinking like a lawyer" that first year students in American law schools must learn. A hallmark of legal imagination, it is also a distinctive feature of U. S. Supreme Court (SCOTUS) oral arguments, "a way for justices to determine the ramifications of choosing one decision over another" and to explore how "policy choices will hold up in slightly different circumstances and factual patterns" [10]. Judges, advocates, law professors, and students practice hypothetical reasoning when they critique a proposed test for deciding a case by posing hypothetical examples. The hypotheticals challenge whether the proposed test is too broad or too narrow, and invite responses to distinguish the hypothetical from the case at hand, modify the test, or abandon the test in favor of another.

In this paper we address some open research questions, including how to model the process of hypothetical reasoning in order to explain its role in legal argument, how to implement the process computationally for purposes of teaching students, and how to evaluate such a model. We present a process model of arguing with hypotheticals and demonstrate how it accounts in a natural way for common features of SCOTUS legal arguments. We explain how, and the extent to which, the process model has been implemented computationally in the LARGO (Legal ARgument Graph Observer) intelligent tutoring system. The program supports students in diagramming SCOTUS oral argument examples in accord with the process model; its feedback on students' diagrammatic reconstructions of the examples enforces the model's expectations.

Finally, we present empirical evidence that features of argument diagrams made with LARGO are correlated with two independent measures related to argumentation ability: standardized test scores that assess ability to evaluate reasoning and arguments and students' number of years in law school. Given the SCOTUS examples and the evidence of diagnostic utility of LARGO diagrams, we conclude that the process model explains some features of legal argument. Finally, we suggest how LARGO diagrams relate to current schemes for representing legal argument diagrammatically.

2. Process Model of Hypothetical Argument

Hypothetical reasoning plays a role in SCOTUS and common law decision making [6; 10; 18], in American legal education [7; 22; 23 pp. 66, 68, 75], in civil law legal reasoning [14, pp. 528-9], in ethical reasoning [9] and in mathematical discovery [11].

The process model of hypothetical argument (PMHA) summarized in Figure 1, an expanded and refined version of the one introduced in [2], was developed to account for frequently observed features of SCOTUS oral arguments. Advocates propose a test (i.e., a general rule) for deciding the case at hand (step 1). The justices then critique the proposed test by posing a hypothetical fact situation and asking how the proposed test would handle it. They do this to probe the advocate's conception of the test and to critique it as too broad (steps 2 and 3) or too narrow (steps 2′ and 3′ below the ellipsis). At step 3 (or 3′), the advocate may: (a) adhere to the proposed test and justify it as correctly deciding the case at hand despite the hypothetical, (b) modify the proposed test to accommodate the hypothetical but still reach the desired result with the case, or (c) abandon the test in favor of another. When responding that the test is not too broad (step 3), justifying the test (3.a) involves analogizing the hypothetical and the case; modifying the test (3.b) involves distinguishing the hypothetical from the case. Responding that the test is not too narrow (step 3′) involves just the reverse: distinguishing in 3′.a and analogizing in 3′.b.

The following examples illustrate three paths through the PMHA. The examples are all drawn from the case of *California v. Carney*, 471 U.S. 386 (1985), where the Court had to decide if police needed a warrant to search a parked motor home. Police had suspected defendant Carney of trading drugs for sex in his motor home located in a downtown San Diego parking lot. After questioning a boy leaving Carney's motor home, agents entered without a search warrant or consent, observed drugs, and arrested Carney. Carney moved to suppress the drug evidence, the State Supreme Court agreed, but the State of California appealed to the U.S. Supreme Court.

SCOTUS oral arguments occur after the parties have submitted briefs but before the Justices decide a case or draft an opinion; each side's advocate has one half hour to press his case before the nine Justices. For the State, an advocate, Mr. Hanoian (Mr. H) argued that the motor home should be treated like an automobile. He proposed a test and defended that it would serve the principles at stake: If the place-to-be-searched has wheels and is self-propelling, then no warrant should be required. This test would prevent the loss of evidence in an emergency situation and would be a "bright line" rule that the police could easily apply. Mr. Carney's advocate, Mr. Homann (Mr. Ho.) posed an alternative test: If the place-to-be-searched has the indicia of a home then a warrant is required. Only such a rule, he argued, would preserve the constitutional principle of autonomy and privacy in ones home.

→ **1. Propose test: For proponent, propose test for deciding the current fact situation (cfs):** Construct a proposed test that leads to a favorable decision in the cfs and is consistent with applicable underlying legal principles/policies and important past cases, and give reasons.

← **2. Pose hypothetical: For interlocutor, pose hypothetical example to probe if proposed test is too *broad*:** Construct a hypothetical example that:

(a) emphasizes some normatively relevant aspect of the cfs and
(b) to which the proposed test applies and assigns the same result as to the cfs, but
(c) where, given legal principles/policies, that result is normatively wrong in the hypothetical.

→ **3. Respond: For proponent, respond to interlocutor's hypothetical example showing test too broad:**

(3.a) Justify the proposed test: Analogize the hypothetical example and the cfs and argue that they both should have the result assigned by the proposed test. *Or*

(3.b) Modify the proposed test: Distinguish the hypothetical example from the cfs, argue that they should have different results and that the proposed test yields the right result in the cfs, and add a condition or limit a concept definition so that the narrowed test still applies to the cfs but does not apply to, or leads to a different result for, the hypothetical example. *Or*

(3.c) Abandon the proposed test and return to (1) (i.e., construct a different proposed test that leads to a favorable decision in the cfs and is consistent with applicable underlying legal principles/policies, important past cases, and hypotheticals…)

• • •

← **2′. Pose hypothetical: For interlocutor, pose hypothetical example to probe if proposed test is too *narrow*:** Construct a hypothetical example that:

(a) emphasizes some normatively relevant aspect of the cfs, and
(b) that normatively should have the same result as the cfs, but
(c) to which the test does not apply or assigns a different result.

→ **3′. Respond: For proponent, respond to hypothetical example showing test too narrow:**

(3′.a) Justify the proposed test: Distinguish the hypothetical and the cfs, arguing that they should not have the same result or that they should have the same result but for different reasons. *Or*

(3′.b) Modify the proposed test: Analogize the hypothetical example to the cfs, conceding that the result should be the same in each and arguing that the proposed test yields the right result in the cfs, and eliminate a condition or expand a concept definition so that the test applies to both the cfs and the hypothetical example and leads to the same result in each. *Or*

(3′.c) Abandon the proposed test and return to (1) (i.e., construct a different proposed test that leads to a favorable decision in the cfs and is consistent with applicable underlying legal principles/policies, important past cases, and hypotheticals…)

Figure 1. Process Model of Hypothetical Argument

Much "action" in oral arguments involves debating about how best to formulate a rule for deciding the case. As in Figure 2, for instance, a Justice attacked Mr. Hanoian's proposed test as too broad. The hypothetical roots the motor home more permanently in a mobile home park with utilities hook-ups. The change in facts may seem innocuous, but it has a profound effect, emphasizing the domicile-like nature of the motor home and deemphasizing the likelihood of the vehicle's fleeing with evidence aboard. In this situation, the need to protect privacy in the home may trump the policy of protecting against evidence loss. Mr. H decided to stick with his test; he analogizes the hypothetical (or "hypo") to the *Carney* motor home, arguing that police cannot know if it has been rooted long or permanently enough not to threaten loss of evidence.

In response to further hypotheticals, Mr. H modified his test as in Figure 3. The wheels requirement was changed to a more general "indicia of mobility". A Justice attacked this as focusing too narrowly on vehicles. A camper's tent pitched next to the motor home would be just as quickly movable, and yet the test would not apply. Arguably, a tent also suggests more clearly that someone is living in it as a home. Mr. H again sticks with his test, distinguishing the tent from the *Carney* motor home, arguing that police can more readily tell if a vehicle is movable than if a tent is lived in.

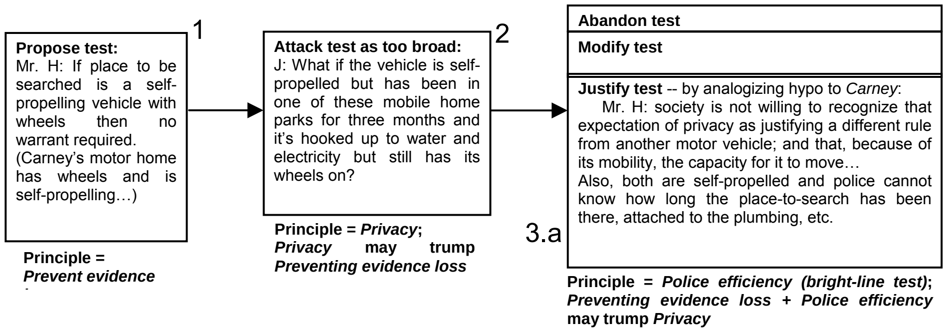

Figure 2: Example of Attacking Test as Too Broad and Justifying Test (#'s keyed to PMHA steps in Fig. 1

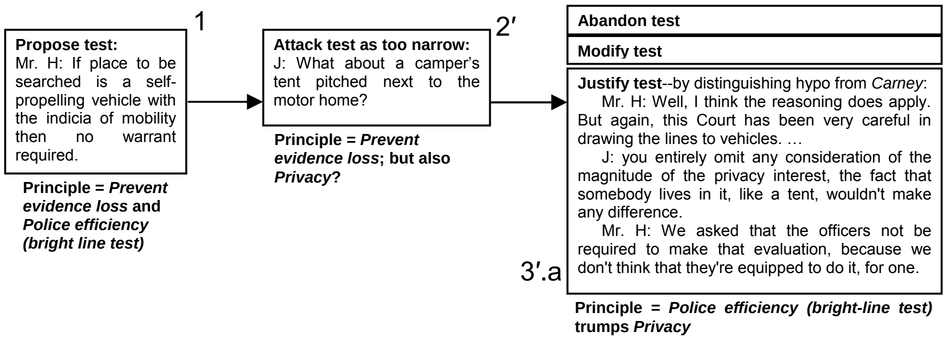

Figure 3: Example of Attacking Test as Too Narrow and Justifying Test

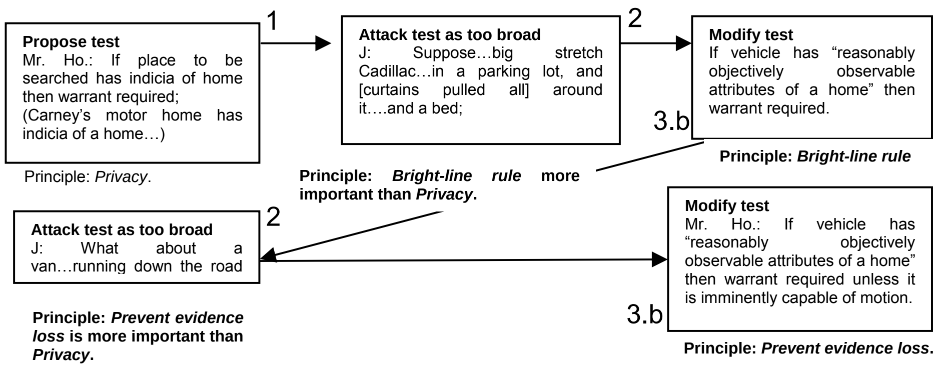

Figure 4: Example of Attacking Test as Too Broad and Modifying Test

Figure 4 shows a different response to an attack that a test is too broad. Here Mr. Ho., the advocate for Carney, twice modifies his test to exclude a hypothetical curtained Cadillac that police could not be sure is a home and a fleeing van.

Although the *Carney* facts invite hypotheticals, and this case has been a focal example in a number of works [5; 9; 15; 20], it is not unique. We have assembled examples of the PMHA from other SCOTUS oral arguments, including cases involving personal jurisdiction, freedom of religion, and contributory copyright infringement. Sometimes the Justices pose hypotheticals, not to critique the test, but simply to explore whether and how it would apply. It may be unclear whether the test applies to the hypothetical, assigns it the same result as the case, or leads to a normatively wrong result. In modeling this more exploratory use of hypotheticals, we can relax certain criteria in step 2 or 2′ of the PMHA, Figure 1, but we do not pursue that here.

The process model of hypothetical argument of Figure 1 adapts patterns of hypothetical reasoning Eisenberg observed in legal opinions to a dialogue between an advocate and a judge [6, p. 100]. It adapts Gewirtz's three common modes of responding to resolve the dissonance created when a proposed test reaches the wrong result in a hypothetical [7, pp. 120f]. It focuses on accommodating the conflicting underlying principles at stake [9, pp. 221-223 226-228]. The model is similar to Lakatos' mathematical reasoning method of proof and refutations [11, p. 50]. SCOTUS oral arguments are real working examples of reasoners' employing local and global hypothetical counterexamples as in the artificial Socratic tutorial dialogue Lakatos reconstructed from centuries-long communications of mathematicians. The PMHA adapts HYPO's 3-ply case-based argument model to a more complex kind of legal reasoning involving underlying principles [1]. In implementing a computational version of the PMHA that actually engages in reasoning with hypotheticals, techniques for broadening or narrowing a legal rule and for reasoning with values would clearly be relevant [3; 4; 21], but they would need to be integrated in a way not yet achieved.

3. LARGO Implementation of Model of Hypothetical Argument

American law students encounter hypothetical reasoning primarily in Socratic classroom dialogs. Students read a small number of cases on a legal issue. In class, the instructor may invite a student to formulate the test that courts appeared to employ in deciding the cases. Like the Justices, the instructor then poses hypotheticals to probe the student's proposed test and how well it accommodates underlying principles and policies across a variety of scenarios. Unlike SCOTUS arguments, there is no official transcript of the classroom exchanges. Students take notes, but the exchanges are fleeting; students are more likely to focus on the resulting rules than on the process that led to them. For this reason, we believe, it would be pedagogically valuable for law students to reconstruct the process of hypothetical argument as revealed in SCOTUS oral arguments of cases relevant to the legal issues students studied.

The LARGO program is intended to help students learn the process of arguing with hypotheticals by diagrammatically reconstructing the arguments in terms of the PMHA. Figure 5 shows the LARGO screen. A SCOTUS oral argument transcript (here from *Carney*) appears in the scrollable pane at the left side of the screen. Students prepare diagrams in the workspace at the right using nodes and links (i.e., elements and relations) drawn from the palette at the bottom left. A student diagrams the argument by selecting the element or relation from the palette, dragging it to the workspace, dropping it there, and connecting it into the developing diagram. Students also link the diagram's elements to passages in the transcript using a text highlighting feature.

LARGO's palette, diagramming-support, and advice-giving all are informed by the PMHA. The palette includes elements for representing the facts of the case for decision, proposed tests, hypotheticals, and five kinds of relations among them: modifying a test, distinguishing or analogizing a hypothetical, a hypothetical's leading to test or modification, and a generic relation. The test element is structured to encourage students to prepare a logical formulation with slots for "if", "then", "and", "unless", and "even though". The diagram in Figure 5 corresponds to the PMHA: A hypothetical led to a proposed test's modification. A second hypothetical with respect to the modified test has been distinguished from the facts of the case.

LARGO implements a computational version of the PMHA, not to make or respond to hypothetical arguments but to help students with their argument diagrams. At any point, a student can select the Advice button at the left and the program responds with three new hints on improving the diagram or reflecting on its significance. The advice tells students: where to look in the transcript for passages that should be represented in the diagram, how to fix up parts of the diagram that appear to be non-standard given the PMHA, and what configurations of diagram elements appear to be worth reflecting about in terms of the model. In producing this advice, the program employs a "graph grammar", a set of rules that enforces the expectations embodied in the PMHA. The rules flag portions of the diagram where the elements and relations miss relevant parts of the text, do not conform to the PMHA, or are complete enough to warrant reflection. The grammar parses the diagram to identify dialectical patterns that trigger feedback; it is applied to a representation of the diagram in graph notation [16]. The grammar rules operationalize a set of concepts for characterizing diagram patterns that indicate if the diagram is consistent with the PMHA.

The classification concepts include indications that elements or relations were not used in the diagram (e.g., No_facts), elements were not linked to the argument text (e.g., Unlinked_test), mistaken relations were used (e.g., a test was related to the case facts by other than a generic relation: Test_facts_relation_specific), or certain patterns are worth reflecting about (e.g., the diagram has at least two tests and a hypothetical: Discuss_hypo_multiple_tests). Thus, while an advocate may have analogized or distinguished a hypothetical and the case facts or modified a proposed test in response to a Justice's hypothetical, a student's diagram may not show these (i.e., No_facts or Isolated_hypo) or it may show uncommon relationships, such as distinguishing a test and the case facts (i.e., Test_facts_relation_specific), or it may show them all and be worth reflecting about (i.e., Discuss_hypo_multiple_tests.)

The graph grammar flags all such mistakes and opportunities and then prioritizes the help in terms of whether it applies to a part of the transcript or the workspace that the student is working in currently, similar advice was given recently, and in which of five localized "phases" the student appears to be in that part of the diagram: (1) orientation, (2) transcript mark-up, (3) diagram creation, (4) analysis, or (5) reflection. LARGO's advice is couched as a recommendation rather than as a declaration that something is incorrect. The program does not have a "definitive" argument representation; an instructor's marked-up transcript only indicates where process-model-related components are located in the text. In addition, a Justice may interrupt the advocate and move on to another topic before the advocate can finish; the diagram will be incomplete according to the model but it accurately reconstructs the argument.

Some dialectical patterns in the students' argument diagrams present "teachable moments", pedagogical opportunities for prompting students to reflect on the policies and principles that inform the hypotheticals and responses. Where a student has analogized or distinguished a hypothetical and the case, LARGO encourages him to explain why this matters (e.g., "Usually, attorneys should give a reason why the distinction matters from a legal viewpoint. For instance, does it matter in terms of the principles and policies underlying the issue? Please enter this in the highlighted distinction relation.") Where a diagram relates one or more tests, hypotheticals, and case facts (i.e, Discuss_hypo_multiple_tests is true), it encourages the student to reflect on the role of the hypothetical vis-à-vis the test, for example: "Does applying the test to the hypothetical represent an acceptable tradeoff of the underlying policies/principles? Did the attorney change his test in response to the hypothetical or not?" or "Do the two tests effect different tradeoffs of the underlying policies/principles?"

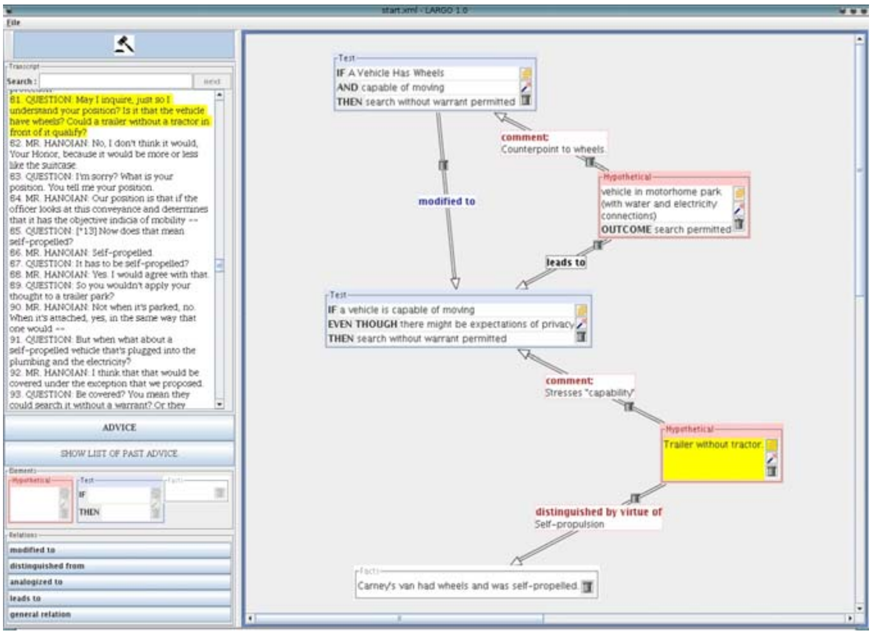

Figure 5: Sample LARGO Diagram of *Carney* Oral Argument

4. Evaluating Process Model of Hypothetical Argument with LARGO

Having developed a multiple-choice test to measure learning and transfer of skills of hypothetical reasoning, we undertook two experiments in fall, 2006 and fall, 2007 to assess how well students learned using LARGO. Both experiments involved first year law students in Legal Process, a required course that introduces students to legal analysis and argumentation. In 2006, 28 paid volunteers were drawn from three sections of Legal Process. In 2007, 70 non-volunteer students from one section of Legal Process participated in the study as a course requirement. In each study, subjects were assigned randomly to either a Diagram condition, using LARGO's graphical support and feedback as described above, or a Text condition that used a text-based word-processing and highlighting environment. In both conditions, students were introduced to the PMHA and, over the course of two (or three in 2007) two-hour sessions, studied SCOTUS oral arguments in two (or three) cases relevant to the course. Their main task was to read the arguments and represent the hypothetical reasoning. As reported in [17] the 2006 evaluation showed promising learning results, especially for volunteer law students with lower LSAT scores, who learned skills of hypothetical reasoning better in the Diagram than in the Text conditions. However, we could not reproduce the learning results with the non-volunteer students in the 2007 study [17]. While use of LARGO's advice correlated with learning, the non-volunteers used the advice much less often than volunteers, who, of course, were self-selected.

Given the disappointing learning results, we undertook a third experiment in spring, 2008, to assess, among other things, whether the process model of hypothetical argument (PMHA) captured something important about real legal argumentation. We began with two assumptions: (1) The Law School Admission Test (LSAT) is a standardized test taken prior to applying for law school "designed to measure skills that are considered essential for success in law school" including "the ability to think critically; and the analysis and evaluation of the reasoning and arguments of others" [12]. (2) A law school education trains students in skills of legal argumentation.

As discussed above, LARGO's diagramming support is based on the PMHA. We hypothesized that if the PMHA captured something important about legal argumentation, then, one would expect, some features of argument diagrams students created with LARGO would reflect differences in their LSAT scores and number of years in law school. To test this hypothesis, we arranged for 25 third year paid volunteer law students to perform the same activities as the Diagram condition in the 2007 experiment, and we then compared the argument diagrams produced by students in the Diagram condition across the three studies (2006, 2007, and 2008) using statistical analysis in relation to LSAT scores and years in law school (i.e., first year (1L) vs. third year (3L)).

We found that the argument diagrams of students of different abilities and experience differed systematically to some extent [13]. With respect to LSAT scores, for the 2007 experiment involving 1Ls, the relations-to-node ratio, a measure of how connected the nodes in the diagrams were to other nodes, correlated positively with students' LSAT scores ($r=.32$, $p<.05$) as did the number of relations ($r=.32$, $p<.05$). We observed a similar trend for the 2006 1Ls' relations-to-node ratio but not for the number of relations. We did not observe a statistically significant correlation between the number of diagram elements and LSAT scores for any of the 1Ls. The 3Ls in 2008 did not evidence any of the above correlations, but, one might expect, LSAT scores taken three years before would have become stale.

With respect to years in law school, across the three experiments, an ANOVA with post-hoc Tukey tests showed that 3Ls' diagrams had significantly ($p<.05$):

(1) more relations (m=12.3) than those of volunteer 1Ls (m=7.9) who produced significantly more than non-volunteer 1Ls (m=5.2);

(2) more elements (i.e., nodes and relations) (m=10.5) than those of 1Ls, and 1L volunteers' diagrams (m=9.6) had significantly more than 1L non-volunteers' (m=7.5), and

(3) larger relations-to-node ratios (avg. 1.14) than 1Ls' diagrams (avg. .82, .67).

We also compared the diagrams in terms of the process-model-based concepts that LARGO uses to assess which phase a student is in for purposes of giving advice. We found that the best predictors of whether a student is a 1L as opposed to a 3L are:

(1) No_facts (Chi-square, c2(8.61,N=51)=1.00, $p < 0.01$, precision=32/51),

(2) Unlinked_test (Chi-square, c2(4.46,N=51)=1.00,$p < 0.05$, precision=32/51),

(3) Test_revision_suggested (Chi-square, c2(12.40,N=51)=1.00, $p < 0.001$, precision=41/51), and

(4) Test_facts_relation_specific (Chi-square, c2(7.44,N=51)=1.00, $p<0.01$, precision=39/51).

1L subjects showed more instances of No-facts (i.e., failure to represent the current case facts in a node in order to analogize or distinguish a hypothetical) and Unlinked-test (i.e., failure to link the test in the diagram to the oral argument text.) 3L's showed more instances of Test_revision_suggested and Test_facts_relation_specific, both of which occur only in LARGO's later advice phases.

The statistical evidence that features of students' LARGO argument diagrams systematically reflect differences in their LSAT scores and number of years in law school tends to confirm the hypothesis; these features plausibly relate to students' aptitude for and experience in making legal arguments. Given that the LSAT is intended to assess a student's ability to analyze and evaluate the reasoning and arguments of others, it makes sense that students with higher LSAT scores produce more connected graphs with more relations. Their reconstructions are more detailed and connected because these students are more likely to be sensitive to the subtler relations and connections revealed as the argument process unfolds. Similarly, the differences between first and third year students relate to the effect of a legal education in training students to "think like lawyers", including inculcating a greater attention to the text, to carefully formulating proposed legal rules, and to drawing inferences by analogizing and distinguishing the facts of the case, hypotheticals, and precedents.

5. Conclusions

We have presented a process model of hypothetical argument and argued that it describes and explains a common feature of legal argument and pedagogy, as illustrated in Supreme Court oral argument examples. We evaluated the model using statistical analysis to relate features of students' model-based diagrammatic reconstructions of the examples to real-world markers of legal argument abilities. The evidence shows that features of the students' argument diagrams are correlated with students' LSAT scores and number of years of law school study, both of which relate to the ability to make and understand legal arguments. This tends to confirm the diagnostic utility of diagrams made according to the PMHA and to support the model as explaining a realistic phenomenon of legal argument. In order to confirm this, we have begun another study in which instructors, blinded as to the diagrams' source, will develop criteria and evaluate the diagrams.

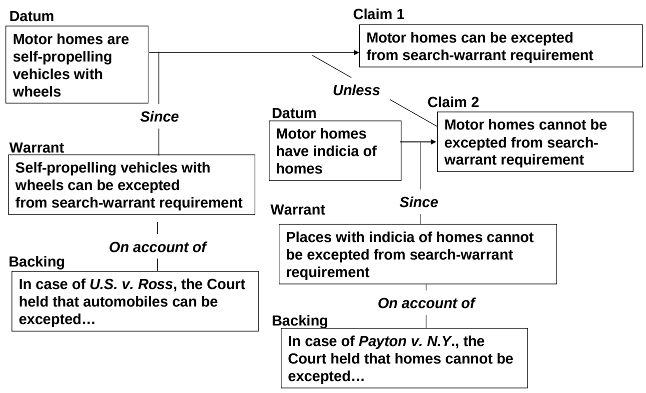

Figure 6: Toulmin Representation of Arguments in *Carney* Case

This evaluation method may help to assess other diagrammatic models of legal argument as in [8; 9; 25]. As in Figure 6, a Toulmin diagram can represent the propositional structure of arguments like those in *Carney*, including rebuttals and case-based or analogical backings [15]. As discussed, however, arguments in the Supreme Court and the legal classroom often involve a strategic process that leads from one version of a warrant to the next. It is as if each warrant in Figure 6 were the culmination of a process of arguing with hypotheticals. A Toulmin-style diagram needs to represent a process, something it ordinarily does not (but see [24; 26]) and complex nested diagrams to represent arguments about the warrant versions [15]. A LARGO diagram does represent a process of generating multiple variations of warrants.

A Carneades version of Figure 6 would also include the status of each statement (e.g., questioned, accepted, rejected) and, given a proof standard, propagate the effects on the claim of changes in status. LARGO diagrams do not propagate any effects, but if they did, they would propagate the effects of changes in a warrant's language and backing. The goal would be to demonstrate how well different formulations and interpretations of warrants fit the facts, past cases, foreseeable hypotheticals and underlying policies and principles. Achieving this goal is our challenge for future work.

6. References

[1] K. Ashley, *Modeling Legal Argument: Reasoning with Cases and Hypotheticals*. The MIT Press (1990).

[2] K. Ashley, N. Pinkwart, C. Lynch, and V. Aleven, Learning by Diagramming Supreme Court Oral Arguments, In *Proc. 11th Int'l Conf. Artificial Intelligence and Law*, 271-275. ACM Press (2007).

[3] K. Atkinson and T. Bench-Capon, Argumentation and Standards of Proof. In *Proc. 11th Int'l Conf. Artificial Intelligence and Law*, 107-116. ACM Press (2007).

[4] T. Bench-Capon and G. Sartor, A Model of Legal Reasoning with Cases Incorporating Theories and Values, *Artificial Intelligence*, 150(1-2):97-143 (2003).

[5] S. Brewer, Exemplary Reasoning: Semantics, Pragmatics, and the Rational Force of Legal Argument by Analogy, *Harvard Law Review*, 109: 923-1028, 937 (1996).

[6] M. Eisenberg, *The Nature of the Common Law*, 99, Harvard U. Press (1988).

[7] P. Gewirtz, The Jurisprudence of Hypotheticals, *Journal of Legal Education*, 32: 120f (1982).

[8] T. Gordon, Visualizing Carneades Argument Graphs. *Law, Probability and Risk* 6: 109 (2007).

[9] S. Hurley, Coherence, Hypothetical Cases, and Precedent, *Oxford J. Legal Studies*, 10: 221, 226ff (1990).

[10] T. Johnson, *Oral Arguments and Decision Making on the U. S. Supreme Court*, 49, SUNY (2004).

[11] I. Lakatos, *Proofs and Refutations*, 50, London: Cambridge University Press (1976).

[12] LSAT.org, http://www.lsat.org/LSAT/about-the-lsat.asp. (visited, August 25, 2008).

[13] C. Lynch, N. Pinkwart, K. Ashley, and V. Aleven, What do argument diagrams tell us about students' aptitude or experience? Workshop on ITSs for Ill-structured Domains, ITS-2008, Montreal (2008).

[14] D. MacCormick and R. Summers (ed.) *Interpreting Precedents* Ashgate/Dartmouth (1997).

[15] S. Newman and C. Marshall. Pushing Toulmin Too Far: Learning From an Argument Representation Scheme. Xerox PARC Tech. Rpt. SSL-92-45 (1992).

[16] N. Pinkwart, K. Ashley, V. Aleven, and C. Lynch, Graph Grammars: an ITS Technology for Diagram Representations. In Proc. 21st Int'l FLAIRS Conf., ITS Track, Coral Gables. May (2008).

[17] N. Pinkwart, C. Lynch, K. Ashley, and V. Aleven, Reevaluating LARGO in the Classroom: Are Diagrams Better than Text for Teaching Argumentation Skills? In *Proc. ITS-08*, 90-100. (2008).

[18] E. Prettyman, Jr., The Supreme Court's Use of Hypothetical Questions at Oral Argument, *Catholic University Law Review*, 33: 555, 556. Symposium on Supreme Court Advocacy (1984).

[19] C. Reed and G. Rowe, Araucaria: Software for Argument Analysis, Diagramming and Representation. *International Journal of AI Tools* 13(4):961-980 (2004).

[20] E. Rissland, "Dimension-based Analysis of Hypotheticals from Supreme Court Oral Argument" In Proc. 2d Int'l Conf. on Artificial Intelligence and Law. 111-120 ACM Press (1989).

[21] E. Rissland and D. Skalak, CABARET: Rule Interpretation in a Hybrid Architecture, *International Journal of Man-Machine Studies*, 34: 839-887 (1991).

[22] R. Stuckey and others, *Best Practices for Legal Education*, 214-215, Clinical Legal Ed. Assoc. (2007).

[23] W. Sullivan, A. Colby, J. Wegner, L. Bond, and L. Shulman, *Educating Lawyers*, 62, 66, 68, 75 (The Carnegie Foundation for the Advancement of Teaching) (2007).

[24] O. Tans, The Fluidity of Warrants: Using the Toulmin Model to Analyze Practical Discourse in D. Hitchcock and B. Verheij (ed.) *Arguing on the Toulmin Model* Springer (2006).

[25] S. Toulmin, *The Uses of Argument*. Cambridge: Cambridge University Press (1958).

[26] B. Verheij, Artificial Argument Assistants for Defeasible Argumentation, *Artificial Intelligence*, 150: 291-324 (2003).

Acknowledgement: This research is supported by NSF Award IIS-0412830.