Interpretive Reasoning with Hypothetical Cases

Kevin D. Ashley

Professor of Law and Intelligent Systems
Senior Research Scientist, Learning Research and Development Center
University of Pittsburgh School of Law
Pittsburgh, PA 15260
Ashley@pitt.edu

Abstract

Reasoning with hypothetical cases helps decision-makers evaluate alternate hypotheses for deciding a case. The hypotheticals demonstrate the sensitivity of a hypothesis to apparently small factual differences that may require different results because they shift the tradeoffs among conflicting underlying principles. By anticipating variations, the decision-maker seeks to formulate as general and robust a hypothesis as possible. This paper presents a model of the role of hypothetical cases in assessing legal hypotheses and illustrates it with examples drawn from a Supreme Court oral argument. It describes the LARGO program, an intelligent tutoring system to help law students learn the model by graphically representing complex argument examples. LARGO analyzes students' graphs and provides feedback to encourage them to reflect on the examples in light of the model.[1]

Introduction

Reasoning with hypothetical cases is useful in interpretive domains, where problems often have no one right answer. Decision-makers consider reasonable arguments for and against alternate hypotheses for deciding the case. A *hypothesis* is a tentative assumption made in order to draw out and test its normative, logical or empirical consequences. A *hypothetical* is an imagined case that involves a hypothesis. The reasoner designs the hypothetical to draw out and test the consequences.

Hypothetical cases help decision-makers evaluate alternate hypotheses for deciding the case. They demonstrate the sensitivity of a hypothesis to apparently small factual differences that actually may require different results because they trigger the application of, or shift the tradeoffs among, conflicting underlying principles. By anticipating new variations, the decision-maker seeks to formulate as general and robust a hypothesis as possible.

Interpretive reasoning with hypothetical cases is closely related to case-based adaptation (Kolodner, 1995, p. 7). While the hypothetical case is not a solution, it is an adaptation that helps investigate proposed rules for deciding the problem in a manner that is consistent with past cases, underlying principles and policies, and anticipated future cases. Hypothetical reasoning is widely employed in such diverse interpretive domains as mathematical discovery (Rissland, 1984), ethical reasoning (Hurley, 1990), and legal reasoning (Eisenberg, 1988). Since hypotheticals play a role in selecting the best among competing explanations of data (i.e., abduction), they are useful in natural science investigation and intelligence analysis (see Josephson, et al., 2003).

In legal reasoning, a hypothesis often takes the form of a proposed test or standard for deciding an issue in a case before a court. An advocate formulates the test or standard based on the relevant statutory or constitutional texts and the interpretations in past cases involving the issue. The advocate asserts that (1) the proposed test or standard is the right standard for the court to apply in deciding the issue, and (2) when applied to the current fact situation (cfs), the standard yields the outcome urged by the advocate. Judges employ hypothetical cases to assess the proposed tests or standards and their application to the case at hand. In oral arguments before the U.S. Supreme Court, the Justices famously pose hypotheticals to probe the legal consequences of adopting and applying an advocate's proposed test. Their hypotheticals explore the meaning of the test, its consistency with relevant legal principles, policies, and past case decisions, its application to the case facts, and its sensitivity to changes in the facts.

From a pedagogical viewpoint, interpretive reasoning with hypotheticals reflects a kind of imaginative and integrative reasoning associated with true substantive mastery of the law. One who poses hypotheticals must be imaginative enough to create a realistic factual scenario that lays bare the proposed test's limitations given the import of the principles and policies of the underlying legal regulations. Similarly, fashioning a response requires an advocate or student to have mastered the principles and polices, and their application to the facts, flexibly enough to distinguish the hypothetical or to modify the proposed rule so that it separates (or unites) the treatment of the hypothetical and current case in a manner that preserves ones argument and protects the interests of ones client. The Supreme Court oral arguments are a unique source of real

[1] This research is sponsored by NSF Award IIS-0412830. Special thanks to Vincent Aleven, co-PI, Niels Pinkwart, Technical University of Clausthal, and Collin Lynch, GRA, University of Pittsburgh Intelligent Systems Program.

world examples of legal Socratic reasoning. They are, however, quite complex.

Our aim is to invent and adapt CBR technology to enable students more effectively to study and learn from these complex, real-world legal examples. In two stages of research, we are attempting computationally to implement a model of interpretive reasoning with hypotheticals. Our first stage program, called LARGO ("Legal ARgument Graph Observer") is an intelligent tutoring system that allows students graphically to represent the important elements of the model as they find them in real legal arguments. We have developed computational tools that assess the completeness of the students' graphical representations and offer assistance in expanding the representations and reflecting about their significance. The second stage is to enable a program to select the next move in a version of interpretive reasoning with hypotheticals in a much simplified card game domain.

Example of Hypothetical Reasoning

An example drawn from the oral argument in a real Supreme Court case, Lynch vs. Donnelly, 465 U.S. 668 (1983), will illustrate both the phenomenon of interpretive reasoning with hypothetical cases and our model of it. This case involved the city of Pawtucket, Rhode Island, which annually erected a Christmas display located in the city's shopping district. The display included such objects as a Santa Claus house, a Christmas tree, a banner reading "Seasons Greetings," and a nativity scene. The crèche had been included in the display for over 40 years. Daniel Donnelly objected to the display and sued Dennis Lynch, the Mayor of Pawtucket. The question in this case was whether the inclusion of the nativity scene in the city's display violated the Establishment Clause of the First Amendment. The example illustrates the proposed test, hypotheticals, and three types of responses apparent in one section of the transcript (with line numbers indicated).

In more detail, the legal advocate in Fig. 1, l. 30 proposes a rule-like test for deciding the problem in a manner favorable to his client. Usually, the advocate offers supportive reasons, such as that the proposed test explains past case decisions or is consistent with, or appropriately reconciles, principles and policies underlying the law. In response, a Justice may pose a hypothetical scenario as in ll. 47, 52, and 57. The hypothetical may simply be a query about the proposed test's meaning, for example, "Would your test apply in these circumstances?" Or, the Justice's hypotheticals (e.g., a lone nativity scene or an Easter cross) may challenge the advocate's proposed test as a kind of counterexample. That is, the hypothetical may imply an argument like, "Isn't the hypothetical disanalogous to the current fact situation? Given the underlying legal principles and policies, shouldn't a suitable test yield different results for the counterexample and for the facts at hand? And, Mr. Advocate, doesn't your test mistakenly treat them the same?" Alternatively, the Justice may imply that her hypothetical is analogous to the current fact situation, that a suitable test should treat them the same in light of the applicable policies or purposes, but that the advocate's proposed test appears to treat them differently.

→ **Proposed test of advocate** (Mr. Mc Mahon for the City of Pawtucket): 30. If the city's creche is displayed in the context of "celebrating a legitimate national holiday and not promoting religious dogma" then the city's display of a creche does not violate the Establishment Clause.

← **Justice's hypothetical:** "47. QUESTION: Mr. McMahon, do you think that for -- a city could display a nativity scene alone without other displays such as Santa Claus and Christmas trees and so forth?"

→ **Advocate's response by arguing that cfs and hypo are analogous:** "50. MR. MC MAHON: That would obviously be less secular than the record shown in this case. However, the most important part of the context, we would submit, is Christmas itself, and any symbol of Christmas, including the creche, is a symbol of what is a national secular holiday.

← **Justice's hypothetical:** "52. QUESTION: Mr. McMahon, what is your answer to Justice O'Connor's question, supposing there was nothing but a creche?

→ **Advocate's response by arguing that cfs and hypo are analogous:** "53. MR. MC MAHON: My answer is, it would still be permissible, Justice, as a symbol of Christmas."

→ **Advocate's fallback response by distinguishing cfs from hypo:** "55. In this case, however, it is so overwhelmed by secular symbolism that the question presented by Justice O'Connor is not before the Court."

→ **Advocate's fallback response by modifying proposed test:** If the city's creche is displayed in the context of the celebration of a legitimate national holiday and if the city is not promoting religious dogma but recognizing a religious element in the holiday *and if the entire government activity includes secular elements in addition to the creche* then the city's display of a creche does not violate the Establishment Clause.

← **Justice's hypothetical:** "57. QUESTION: Mr. McMahon, could the city display a cross for the celebration of Easter under your view?"

→ **Advocate's response by distinguishing cfs from hypo:** "58. MR. MC MAHON: Justice, I think that, first of all, Easter is not a recognized public holiday, and the association of a cross with a specifically religious holiday might well implicate the promotion of religion."

Fig. 1: Example of hypothetical reasoning in argument

In order to maintain the credibility of his argument, the advocate has to rebut or otherwise reply to such a challenge. He may attempt, as in ll. 50 and 53, to save his proposed test by arguing that the supposedly disanalogous counterexample is really analogous, disputing that a suitable rule applied to the counterexample should yield a different result. Or, he may distinguish the hypothetical example from the current facts as in his fallback response, l. 55, and/or modify the proposed test so that it behaves like a suitable rule or does not apply to the hypothetical at

all. Or, he may concede the Justice's point and abandon the proposed test in favor of another approach.

Model of Hypothetical Reasoning

In more schematic form, interpretive reasoning with hypotheticals can be captured in a heuristic 3-ply model, Fig. 2, intended to guide an intelligent agent in analyzing a case. This heuristic process could guide the interactions of an advocate and a judge in an oral argument, or the reflections of a reasoner like an advocate who is trying to prepare an argument by anticipating possible responses.

→ **Point:** For *proponent*, propose a test and argument for deciding the cfs.
Examine past cases and their decision rules and see if a past rule applied to the cfs arguably leads to a favorable decision. If so, use that rule as a proposed test for deciding the cfs and give reasons. If not, construct a proposed test that, when applied to the cfs, leads to a favorable decision and is consistent with the results of some important past cases, and give reasons.

← **Response:** For *respondent*, pose a hypothetical case as counterexample and argument. Inspect *proponent's* test and argument in light of past cases/rules that arguably lead to the opposite conclusion. Construct a hypothetical counterexample to the proposed test, such that the counterexample is:
- analogous to [disanalogous from] the cfs (i.e., a suitable test when applied to the counterexample should yield the same [a different] result) and yet the proposed test when applied to the counterexample leads to a different [the same] result, and give reasons.

→ **Recovery:** For *proponent*, rebut or otherwise reply to *respondent's* counterexample:
- Save the proposed test by disputing that a suitable rule applied to the counterexample should yield the same [a different] result (i.e., show that the supposedly analogous counterexample is really disanalogous [analogous]). Or
- Modify the proposed test so that it is behaves like a suitable rule or does not apply to the counterexample (i.e., remove [add] a condition (or expand [limit] a concept definition such that the modified rule applies to the counterexample and yields the same result, applies to the counterexample and yields a different result or no longer applies to the counterexample, as appropriate.) Or
- Abandon the proposed test.

Fig. 2: Model of hypothetical reasoning

The heuristic process is an attempt to adapt HYPO's 3-ply arguments (Ashley, 1988; 1990) to a more complex kind of legal argument, one that makes explicit a proposed test for deciding the cfs, a test that is phrased in terms of abstract legal concepts. It involves posing hypothetical counterexamples to the tests and making arguments that a proposed test should or should not apply to the counterexamples. It also requires an ability to modify the proposed test as a way of recovering from the response. It is adapted from Lakatos's three heuristic rules for proof and refutations which employ global and local counterexamples to critique mathematical conjectures (Lakatos, 1976, p.50).

Computational Approaches

In AI machine learning research, (Hayes-Roth, R., 1983) and (Pease, Colton, et al., 2002) describe programs that perform Lakatos-style reasoning with counterexamples in card game play (i.e., Hearts) and with number theory concepts, respectively. It appears, however, that no CBR or AI and Law program has modeled computationally all of the behavior described in the above model of interpretive reasoning with hypotheticals. There have been a number of notable efforts. (McCarty & Sridharan, 1981) developed a program that modeled an argument in one Supreme Court tax case, including reasoning with rule concepts, past cases, and an accepted hypothetical. Our goal is to generalize such a model to account for more aspects of the 3-Ply interpretive process model so that it can apply to a wider range of cases and arguments. HYPO (Ashley, 1988; 1990) posed hypothetical variations of problem situations to strengthen/weaken arguments; it represented cases in terms of factual dimensions and used five heuristics to pose hypotheticals by modifying cases along dimensions. It modeled a number of argument moves relevant here such as analogizing, distinguishing, and posing counterexamples, but not in a dialogical context in which hypothetical counterexamples are used to put pressure on proposed tests and their terms. (Ashley, 1988; 1990) interprets a number of Supreme Court oral arguments in terms of a dimensional model and the five heuristics; (Rissland, 1989) applied this approach to interpreting the oral argument in a warrantless search case. These interpretations did not focus on modeling a number of phenomena of interest here: proposing tests, using hypotheticals to challenge proposed tests, and responding to such challenges. Similarly, CATO (Aleven, 1997; 2003) engaged students in making case-based legal arguments analogizing and distinguishing cases in terms of dimension-like factors and providing reasons that invoked underlying legal issues, but not posing hypotheticals to assess tests. (Bench-Capon & Sartor, 2003; Chorley & Bench-Capon, 2005) present a model and program that represent cases as factors whose significance is based on abstract values, and model arguments about rule-making, that reflect preferences among conflicting values. While relevant to various aspects of the 3-Ply interpretive process model, these do not model the entire behavior.

LARGO Intelligent Tutoring System

In our LARGO system (Pinkwart, et al. 2006) students study transcripts of U.S. Supreme Court oral arguments

and create graphical representations of the argument flow as advocates propose tests for deciding the case, Justices pose hypotheticals to challenge the consistency of the test with legal precedents and principles, and advocates respond to those challenges.

From a pedagogical viewpoint, these oral argument examples are worth studying. While it is important for law students to develop a legal imagination and skills of interpretive reasoning with hypotheticals, these skills are hard for law professors to teach. Ideally, students learn these skills from participating in Socratic questioning in class. It is difficult, however, for law professors to perform hypothetical reasoning consistently well enough for students to understand the method and the deeper legal connections it reveals, especially since students in large law school classes often are bystanders and not active participants. The examples are not written, they appear only briefly in classroom discussion before the professor moves on, and the class rarely steps back to reflect explicitly on the argument moves but only on the content they reveal about the underlying legal domain. A program that engages students in reflecting upon expert examples of hypothetical reasoning would be very helpful.

The LARGO system analyzes the law students' developing argument diagrams for different types of potential weaknesses. As an output, it prompts students to remediate the apparently weak parts of the argument representation; subsequent prompts invite students to reflect on the significance to the argument's merits of the relationships among tests, hypotheticals, and responses. The LARGO system was built using the Cool Modes framework (Pinkwart, 2005) as a technical base.

A sample graphical representation, prepared by the author using LARGO, of the portion of the oral argument in the Lynch v. Donnelly case discussed above, is shown in Figure 3. The transcript of the oral argument appears on the left side of the screen. Below that appears a palette with the basic elements a student can use to construct a graphical representation of the argumentation transcript. The basic elements include the current fact situation, proposed tests (and modifications), hypotheticals, and various relations among these (e.g., test modification, distinction of hypothetical, hypothetical leading to test change, and general relation). An "Advice" button leads to LARGO's feedback given the particular state of the argument graph that the student has constructed in the workspace on the right. Students create the graphs using simple drag and drop mechanisms. They can easily link particular diagram elements to specific passages in the transcript using a text highlighting feature. The diagram contains two test versions, two hypotheticals, one element representing the current fact situation (cfs), and several relationships between these elements. For example, the hypothetical concerning the "lone nativity scene" is distinguished as less secular than the cfs and leads to the modified test version. The "cross of Easter" hypo (highlighted) is distinguished from the cfs as "not a recognized national holiday".

By means of a specially designed graph grammar, LARGO detects potential weaknesses in the structure and context of the graphical representations (Pinkwart, et al. 2006). Structural weaknesses involve portions of the graph that do not evidence the relations among elements that correspond to the model of interpretive reasoning with hypotheticals (Fig. 2). For instance, advocates often distinguish a hypothetical from the cfs or analogize the two; one version of a test is related to another and may have been proposed in response to a hypothetical, but these relationships may not be present in the student's graph. On the other hand, ordinarily it does not make sense to distinguish a hypothetical from a test. Context weaknesses involve the absence in a student's graph of elements corresponding to passages that the professor's marked-up version of the argument transcript indicates contain important elements. For example, the student may have missed a proposed test or a hypothetical.

LARGO's graph grammar has heuristic rules to detect both types of anomalies; they are heuristics because in the ill-defined domain of legal argument, one cannot be certain that the missing or mistaken elements or links are necessarily erroneous, or exactly where they occur in the transcript. In Fig. 3, for instance, the Advice button will lead to a suggestion to review the unlikely "distinguished" link between the Easter cross hypothetical and the first test. If this hypothetical were not evident in the graph, the advice would suggest that the student review line 57.

LARGO employs a collaborative filtering approach to detect and deal with certain content weaknesses in argument representations. This enables it to ask whether a student's free-text description of an aspect of a legal argument, such as his restatement of a proposed test, is adequate without employing natural language processing. The collaborative filtering is based on students' selections among different sample formulations of a test that other students or the professor have categorized as more or less effective (Pinkwart, et al. 2006).

LARGO has a heuristic mechanism for identifying the phase that a student appears to be in with respect to the graph representation task (i.e., (A) system orientation; (B) text reading, (C) highlighting and linking; (D) connection of diagram elements; (E) diagram analysis and editing; or (F) reflection). It uses this phase information to select from all of the possible topics of advice (e.g., potential weaknesses) that apply at any given time, the best ones to raise. Once the student's graph is fairly complete, the advice urges the student to reflect on the significance of various elements and their relationships to the merits of the argument. For instance, in the graph of Fig. 3, once LARGO detects the student is in phase (F), the Advice function would highlight the "lone nativity scene" hypo and the leads-to link to the revised test, as well as the distinction from the cfs, and offer a hint for reflection: "Please elaborate on the impact of the hypothetical in the text box below. Did the attorney change his test in response to the hypothetical or not?..."

In an empirical study, we have assessed the utility of the LARGO approach and the effectiveness of its feedback in a first semester Legal Process class using the oral arguments from U.S. Supreme Court cases involving personal jurisdiction and contributory copyright infringement (as a transfer problem). We have represented oral arguments across four unrelated legal domains (including warrantless search of a motor home), increasing our confidence that the model of hypothetical reasoning and its representation with LARGO are quite general. Since the approach does not require a highly detailed mark-up of or deep system-side knowledge about specific legal arguments, only modest effort has been needed in order to use the system with a new transcript in a new legal domain. We compared any learning gains of students using LARGO's graphical argument representation and feedback with those of a control group reading the same oral arguments using a traditional text-based word-processing environment that supports highlighting and note-taking. Students in both groups were asked to identify and relate the elements of the same dialectical patterns of hypothetical reasoning. Using objective pre- and post-tests, we showed that students with lower LSAT scores learned argumentation skills, as measured in a near transfer problem, significantly better using LARGO than comparable students using the text-based tool. In addition, in evaluating hypotheticals with respect to a test, LARGO subjects with lower or medium LSAT scores significantly outperformed comparable subjects in the control condition. Both results support the pedagogical value of the LARGO approach.

Toward Hypothetical Reasoning in a Game

The LARGO program does not perform case-based reasoning per se; it helps students understand and reflect upon textual examples of an important kind of CBR process. It would be delightful if a program could also engage students in interpretive reasoning more directly by posing hypotheticals or responding to them according to the 3-ply model in Fig. 2. Creating fact representation schema, however, even for only the four legal domains for which we have collected oral arguments, is a substantial undertaking. Before tackling it, we hope to gain some assurance that we can develop the inference and control mechanisms that would enable a program to engage in the 3-ply argument behavior.

We are designing a simplified 2-player card game domain in which to experiment with these mechanisms. Each "case" is a triplet of cards (e.g., 3□, Q□, J□); the suit of the third card (□) is the outcome. The object of the advocate-player is to make a point by proposing a rule-like hypothesis to explain the outcome in terms of the other two cards in a manner as consistent as possible with a small set of procedural and substantive principles for the domain, such as, Coverage (i.e., decision is consistent with as many past "case" decisions -- previously "decided" card triplets-- as possible) or "Like makes right: similarity in cards should be rewarded with □ or □." The goal of the judge-player is to respond by posing a hypothetical "case" (i.e., card triplet drawn from a hand) with which to challenge the proposed rule as per the 3-ply model. The advocate-player must then decide how to recover, perhaps by redefining one of the concepts in the proposed rule or modifying it in some other way.

Given the simplified domain and cases, the inferences needed to (re)define concepts for the hypotheses, apply them to the existing cases, or use them to fashion relevant hypotheticals that relate the proposed rules to the underlying principles are simpler (but by no means trivial.) Our goal is for the program to take either role in playing the game with students. We hope that the resulting argument moves with proposed rules, hypotheticals and responses are intuitively intelligible to human "judges" and, when blinded as to the source of the moves, that the judges will be unable to tell whether the program or a student player has generated a good move.

Conclusion

Reasoning with hypotheticals is an important variation of case-based reasoning employed in a variety of interpretive domains, such as law, ethics, and mathematical discovery. It involves proposing hypotheses and posing hypothetical cases in order to clarify and challenge them. This systematic study of interpretive reasoning with hypotheticals will model aspects of imaginative and creative case variation and provide insights into adaptation in CBR. Hypothetical reasoning is like case-based adaptation, except that the modified cases are not so much solutions to a problem as probes in order to evaluate a hypothesis and explore modifications of the hypothesis that might better serve as solutions. In probing hypotheses by inventing new realistic cases and variations, hypothetical reasoning explores deeper connections among facts, proposed hypotheses or decision rules, and principles and policies underlying the domain.

References

Aleven, V. (1997) *Teaching Case-Based Argumentation Through a Model and Examples*, Ph.D., U. Pittsburgh.

Aleven, V. (2003) "Using Background Knowledge in Case-Based Legal Reasoning" *Artificial Intelligence* 150(1-2):183-238.

Ashley, K. (1990) *Modeling Legal Argument: Reasoning with Cases and Hypotheticals* (The MIT Press). based on (1988) Ph.D.. Tech. Rep. No. 88-01 COINS, U. Mass.

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

Chorley, A. and Bench-Capon, T. (2005) AGATHA: Automated Construction of Case Law Theories through

Heuristic Search, Proc. Tenth Int'l Conf. Artificial Intelligence and Law 45-54 ACM Press.
Eisenberg, M.A. (1988) *The Nature of the Common Law* Harvard U. Press.
Hayes-Roth, R. (1983) "Using proofs and refutations to learn from experience",in *Machine Learning: An Artificial Intelligence Approach*, eds., R. S. Michalski, et al., 221-240, Tioga, Palo Alto.
Hurley S. L.,(1990) Coherence, Hypothetical Cases, and Precedent, 10 *Oxford J. of Legal Studies* 221, 230-234.
Josephson, J., Chandrasekaran, B. and Carroll, M. (2003) "Toward a Generic Architecture for Information Fusion". Proc. Collaborative Technology Alliances Conference on Advanced Decision Architectures. pp. 111-116.
Kolodner, J. (1993) *Case-Based Reasoning*. Morgan Kaufmann: San Mateo.
Lakatos, I. (1976) *Proofs and Refutations*. London: Cambridge U. Press.
McCarty, L.T. and Sridharan, N. S. (1981) The Representation of an Evolving System of Legal Concepts: II. Prototypes and Deformations (LRP-TR-11. Lab. for CS Res. Rutgers U.).
Pease, A., Colton, S., Smaill, A. and Lee, J. (2002) Lakatos and Machine Creativity ECAI 2002.
Pinkwart, N. 2005. Collaborative Modeling in Graph Based Environments. Berlin, dissertation.de Verlag.
Pinkwart, N., Aleven, V., Ashley, K. and Lynch, C. (2006) Toward Legal Argument Instruction with Graph Grammars and Collaborative Filtering Techniques. Proc., 8th Int'l Conf. Intell. Tut. Sys. Springer.
Rissland, E. L. 1984. "The Ubiquitous Dialectic," in T. O'Shea (Ed.) *Proc. 6th European Conf. on Artificial Intelligence Pisa, Italy.* 367-72 North-Holland.
Rissland, E.L. (1989) "Dimension-based Analysis of Hypotheticals from Supreme Court Oral Argument." *Proc.2d Int'l Conf. AI and Law*, pp. 111, Vanc. BC .

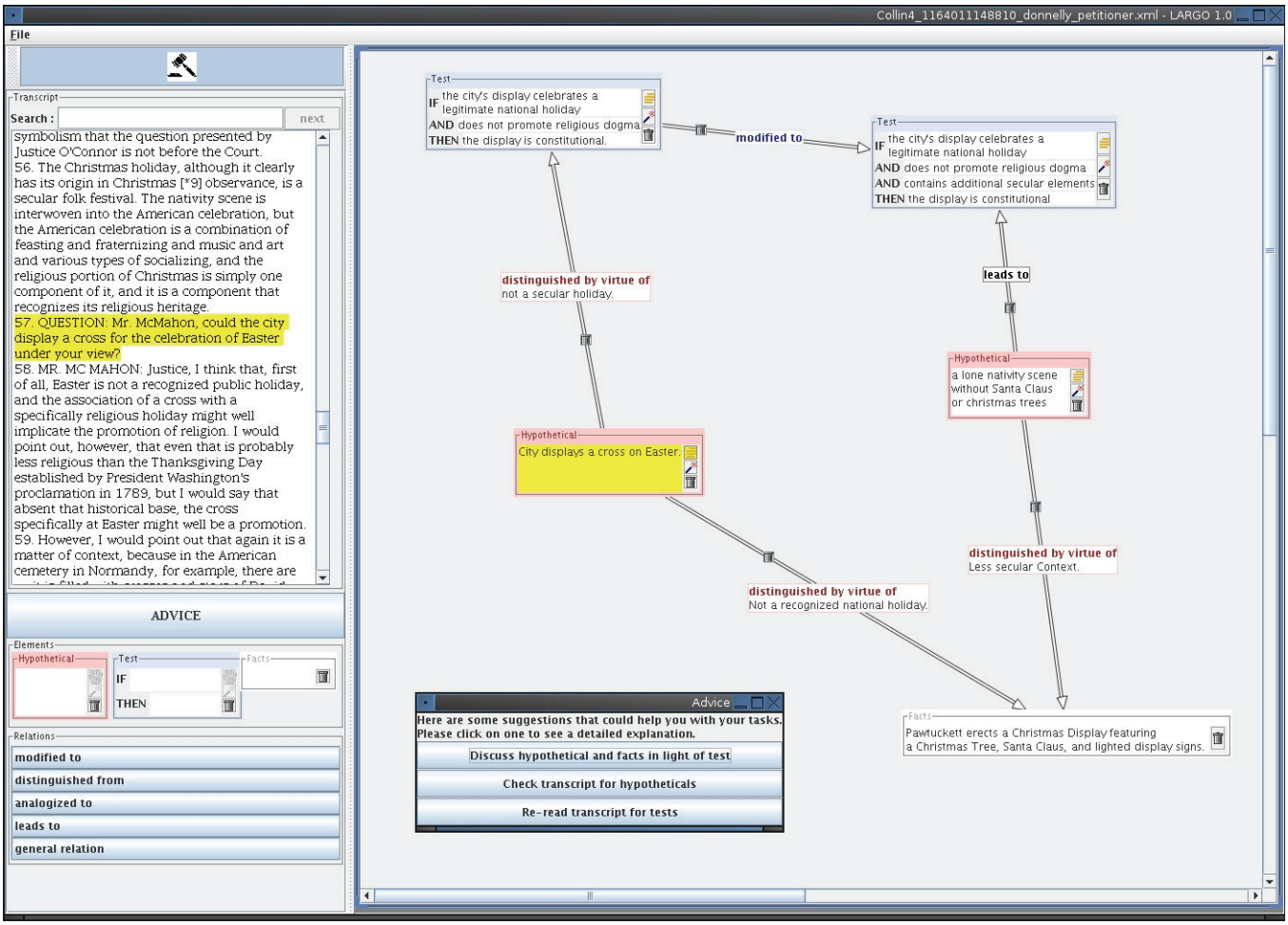

Fig. 3: Example of a graphical argument model in LARGO