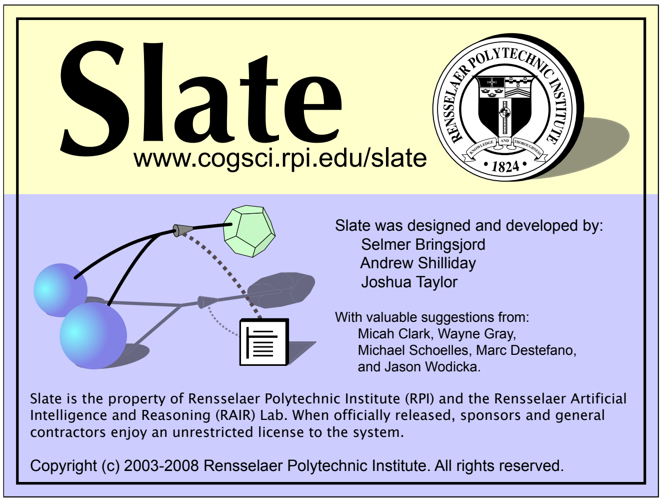
Slate: An Argument-Centered Intelligent Assistant to Human Reasoners

Selmer Bringsjord and **Joshua Taylor** and **Andrew Shilliday**
and **Micah Clark** and **Konstantine Arkoudas**[1]



Slate
www.cogsci.rpi.edu/slate

Slate was designed and developed by:
Selmer Bringsjord
Andrew Shilliday
Joshua Taylor

With valuable suggestions from:
Micah Clark, Wayne Gray,
Michael Schoelles, Marc Destefano,
and Jason Wodicka.

Slate is the property of Rensselaer Polytechnic Institute (RPI) and the Rensselaer Artificial Intelligence and Reasoning (RAIR) Lab. When officially released, sponsors and general contractors enjoy an unrestricted license to the system.

Abstract. We describe Slate, a logic-based, robust interactive reasoning system that allows human "pilots" to harness an ensemble of intelligent agents in order to construct, test, and express various sorts of natural argumentation. Slate empowers students and professionals in the business of producing argumentation, e.g., mathematicians, logicians, intelligence analysts, designers and producers of standardized reasoning tests. We demonstrate Slate in several examples, describe some distinctive features of the system (e.g., reading and generating natural language, immunizing human reasoners from "logical illusions"), present Slate's theoretical underpinnings, and note upcoming refinements.

1 INTRODUCTION

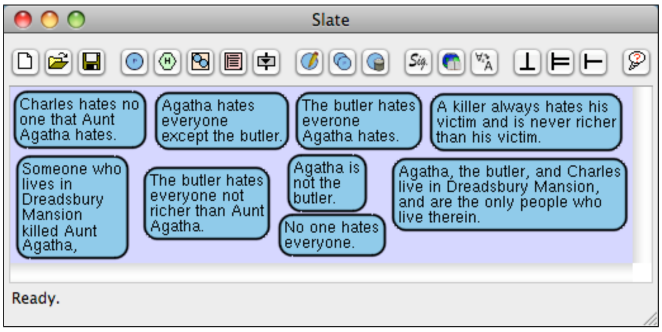
Slate is a robust interactive reasoning system. It allows the human "pilot" to harness an ensemble of intelligent agents in order to construct, test, and express natural argumentation of various sorts. Slate is designed to empower students and professionals in the business of producing argumentation, e.g., mathematicians, logicians, intelligence analysts, designers and producers of standardized reasoning tests, and so on. While other ways of pursuing AI may well be preferable in certain contexts, faced with the challenge of having to engineer a system like Slate, a logic-based approach [9, 10, 18, 31, 13] seemed to us ideal, and perhaps the power of Slate even at this point (version 3) confirms the efficacy of this approach. In addition, there is of course a longstanding symbiosis between argumentation and logic revealed in contemporary essays on argumentation [48]. In this paper, we summarize Slate through several examples, describe some distinctive features of the system (e.g., its capacity to read and generate natural language, and to provide human reasoners with apparent immunity from so-called "logical illusions"), say a bit about Slate's theoretical underpinnings, and note upcoming refinements.

2 A SIMPLE EXAMPLE

We begin by following a fictitious user, Ulric, as he uses Slate to solve a short logic puzzle, the *Dreadsbury Mansion Mystery* [34]:[2]

> Someone who lives in Dreadsbury Mansion killed Aunt Agatha. Agatha, the butler, and Charles live in Dreadsbury Mansion, and are the only people who live therein. A killer always hates his victim, and is never richer than his victim. Charles hates no one that Aunt Agatha hates. Agatha hates everyone except the butler. The butler hates everyone not richer than Aunt Agatha. The butler hates everyone Agatha hates. No one hates everyone. Agatha is not the butler. *Who killed Agatha?*

Information can enter Slate in a number of formats, e.g., as formulae in many-sorted logic (MSL), or as sentences in a logically-controlled English (§4.2). Information can also be imported from external repositories such as databases or the Semantic Web (§4.5). Ulric examines the Dreadsbury Mansion Mystery facts displayed in Slate's workspace (Figure 1).



Charles hates no one that Aunt Agatha hates.

Agatha hates everyone except the butler.

The butler hates everone Agatha hates.

A killer always hates his victim and is never richer than his victim.

Someone who lives in Dreadsbury Mansion killed Aunt Agatha,

The butler hates everyone not richer than Aunt Agatha.

Agatha is not the butler.

No one hates everyone.

Agatha, the butler, and Charles live in Dreadsbury Mansion, and are the only people who live therein.

Ready.

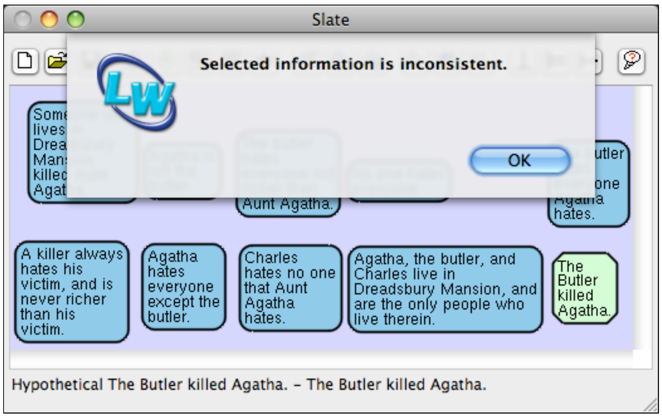
Figure 1. The Dreadsbury Mansion Mystery facts represented in Slate.

A fan of murder mysteries, he considers whether conventional wisdom might hold true, i.e., that the butler did it. Ulric adds the hypothetical to Slate's workspace and asks Slate to check whether the hypothesis is consistent with the other propositions. Slate quickly reports an inconsistency (Figure 2).

---

[1] Rensselaer Polytechnic Institute (RPI), USA, email: {selmer, tayloj, shilla, clarkm5, arkouk}@rpi.edu

[2] This puzzle is of a type typically used to challenge humans (e.g., students in introductory logic courses) and machines (e.g., automated theorem provers).



Selected information is inconsistent.

Hypothetical The Butler killed Agatha. – The Butler killed Agatha.

Figure 2. A murderous butler is inconsistent with the premises.

At this point, Ulric suspects that Charles must have killed Agatha, although he recognizes that he hasn't envisioned a detailed argument. He adds to the workspace the hypothetical that Charles killed Agatha, and asks Slate if it can prove Charles' guilt from the given facts.

Slate's response surprises Ulric: Not only was Slate unable prove Charles' guilt, but the system generated a countermodel! Of course, Ulric realizes, the existence of a countermodel means that Charles isn't implicated *deductively*, but there still might be, for example, an abductive or inductive indictment. Ulric decides to examine the countermodel that Slate found (Figure 3).

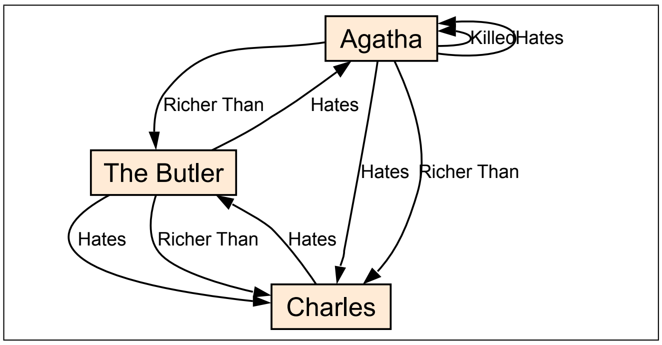
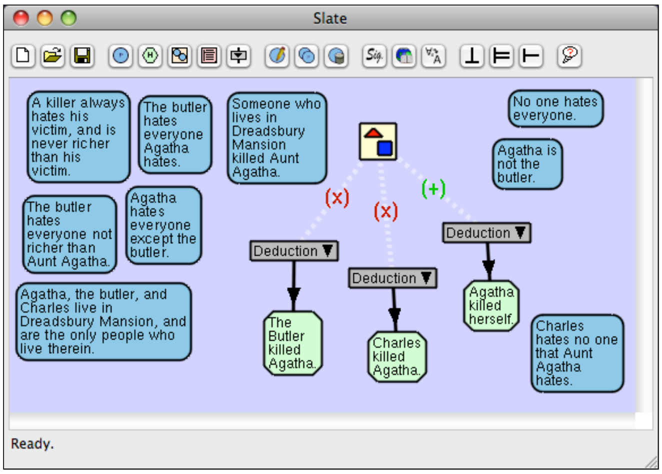


Figure 3. The countermodel casts reasonable doubt on Charles' guilt.

Slate depicts the countermodel as a directed graph, where binary relations between people are denoted by labeled arcs between them. For instance, that Charles hates the Butler is shown by the arc from *Charles* to *The Butler* labeled *Hates*.

What Ulric finds interesting, though, is not that the Butler hates Charles, but that in the countermodel, Agatha killed herself, i.e., it was a suicide. Ulric knows that when Slate finds a countermodel, the facts expressed therein aren't necessarily entailed by the selected propositions, but rather are *mutually consistent*. Until this point, he had tacitly presumed that either Charles or the Butler killed Agatha, and now he realizes that this presumption isn't yet warranted. Ulric decides to investigate the hypothesis that Agatha killed herself, based on the logical possibility that she did, and asks Slate if the new hypothetical can be proven (Figure 4).

Slate responds that there *is* a deductive proof that Agatha committed suicide. At this point, Ulric either examines the natural-deduction style proof that Slate found, or begins his own investigation. Eventually he builds an argument at the level of detail he desires (Figure 5). This argument is mechanically certified by Slate, indicating the logical validity of each inference made therein.



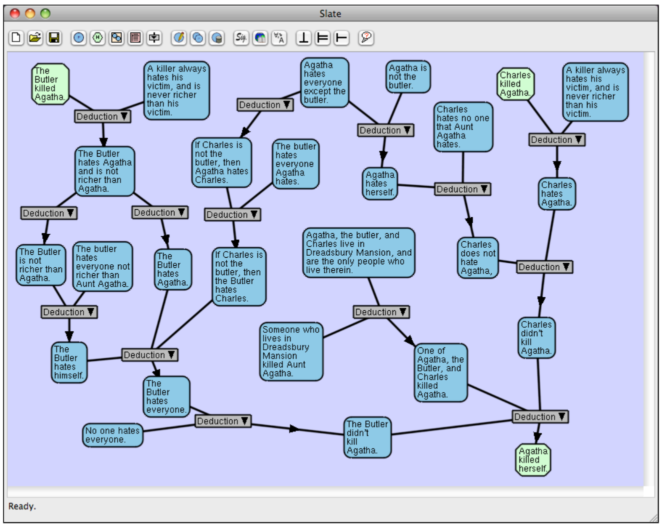
Ready.

Figure 4. The model suggests the hypothesis of Agatha's suicide.

3 THEORETICAL FOUNDATIONS

Slate is based on a robust, multi-faceted theory of heterogeneous human and machine reasoning—a theory that affirms the importance of deductive, inductive, abductive, analogical, and visual reasoning; arguments and counter-arguments; proofs and disproofs; models and counter-models; and strength factors (in the tradition of Chisholm [15] and Pollock [37]) that force explicit declarations of reliability in source and provenance information. This theory should in no way be confused (let alone be conflated) with limited, prior theories of argumentation and argument mapping, e.g., those based on Toulmin's *The Uses of Argument* [47]. As an immediate corollary, note that Slate is radically different than all software systems based on such prior theories. From the standpoint of education and training, Slate is based on a neo-Piagetian view of the development of bias-free human reasoning, according to which, given sufficient training, neuro-biologically normal humans can reason (deductively, and in other modes as well) in normatively correct, bias-free fashion [8, 39].

Slate's purpose is to facilitate and amplify its users' "System 2"



Ready.

Figure 5. A detailed deductive argument proving Agatha's suicide.

cognitive abilities. As Stanovich & West [45] explain, evidence accumulated over a number of decades strongly supports the view that there are two cognitive systems at play in the human mind: “System 1” and “System 2.” Reasoning performed on the basis of System 1 cognition is bound to concrete contexts and is prone to error; reasoning on the basis of System 2 cognition “abstracts complex situations into canonical representations that are stripped of context” [45], and when such reasoning is mastered, the human is armed with powerful techniques that can be used to handle increasingly complex challenges that the modern, symbol-based marketplace presents.

Of course, if Slate is to amplify human cognitive abilities, it must embrace the diverse mechanisms of human reasoning. Humans apparently sometimes reason via “mental logic” [40, 6] (the manipulation of purely linguistic entities, e.g., logical formulae), and sometimes by imagining and manipulating non-linguistic “mental models” [25, 41] of possible situations. Further, as brain studies now increasingly confirm (see e.g. [19, 14]), they also sometimes reason in a fashion involving mental logic, mental models, and *meta-reasoning* over the structures and inference rules of the two prior theories (a process that cannot be captured by either prior theory). This kind of heterogeneous reasoning (explored, e.g., in the theory of *mental meta-logic* [50, 51, 53, 52]), is the impetus for State’s use of proof- and model-theoretic argumentation (and soon, argumentation unifying symbolic and visual information, see §7).[3]

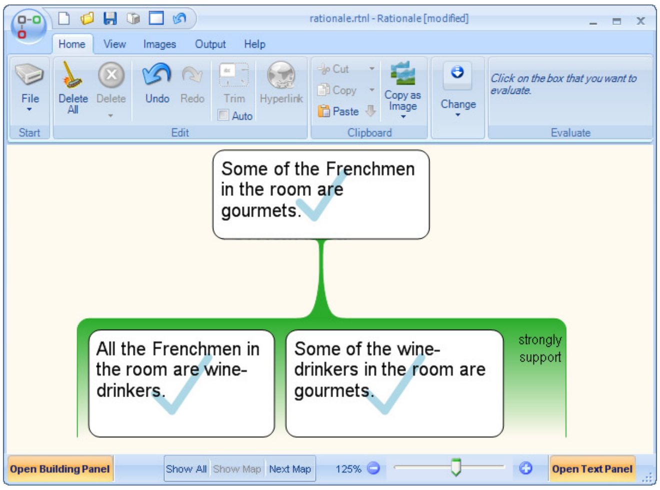
Slate is an interactive system, and must not only draw correct conclusions according to these aforementioned theories, but must also hold to cognitively plausible theories of inference. For example, epistemic theories are, by and large, built on logics of possibility and necessity and interpreted according to Kripke semantics. Undoubtedly, a person’s response, when asked why they hold some particular belief, will not include a description of an accessibility relation partitioning all possible worlds, but rather an articulation of some argument for their belief—perhaps even one with explicit assignments of strength to evidence and inferences. Given that Slate must be able to evaluate such justification, its epistemic theory must be cognitively plausible, based on a heterogeneous system of defeasible argumentation, where validity, veracity, and inferential and evidentiary strength are the basis for justified belief and knowledge.

4 COMPONENTS AND FEATURES

4.1 Visual Interface

It may be surprising that Slate, a system ultimately grounded in formal logic, is inherently visual; logics, with exceedingly rare exceptions, have been non-visual since the first one arrived on the scene 300BC, courtesy of Aristotle. But Slate is a system for argumentation—specifically, a system to *assist* in argumentation—and has therefore adopted a representation that subsumes and exceeds argument-mapping technologies, e.g., Toulmin’s work [47][4] and *Rationale*[5] (Figure 6). The basic construct of the Slate workspace is a proposition or hypothesis, representing an individual fact or statement. An argument in Slate is a tree-like structure with the ultimate conclusion as its root and initial premises at the leaves. Non-leaf nodes within arguments are supported by an inferential link connecting the node (conclusion) to its children (premises).[6] Each inference can indicate the mode of reasoning under which the conclusion is supported by the premises. Reasoners may construct sub-arguments, in which new hypotheticals are introduced and temporarily supposed true.

Of course, humans also reason visually about models (i.e., possible states of the world), and process diagrammatic information such as maps, charts, and graphs. Slate incorporates models and countermodels into arguments. Presently these elements are visualized as digraphs (Figure 3), but more sophisticated visual representations (such as the aforementioned maps, diagrams, charts, plots) are planned for the future, using “visual-logics,” such as Vivid.[7]



Click on the box that you want to evaluate.

Some of the Frenchmen in the room are gourmets.

All the Frenchmen in the room are wine-drinkers.

Some of the wine-drinkers in the room are gourmets.

Figure 6. A simple argument map in Rationale.

4.2 Linguistic Interface

Though there are many benefits to working in a formal logic, most human reasoning and argumentation is performed at the linguistic level. Information is usually represented textually, arguments are presented in (perhaps structured) prose, and reasoning is justified in natural language. Many of Slate’s users do not have mastery of formal logic, and so Slate accepts information expressed in *logically-controlled English*, a subset of English that can be unambiguously translated into a formal logic. The process by which this occurs has three phases [7] (Figure 7):

Phase 1: English texts are rephrased in logically-controlled English. Slate makes use of ACE [17, 22] and CELT [32, 33], each of which is a logically-controlled English with a fixed, definite clause grammar and allows user-defined vocabularies.[8]

---

[3] For in-depth discussion of the aforementioned psychological theories, and their relation to logical systems, see [13]. Please note that proponents of these psychological theories of human reasoning generally view them as entirely incompatible. From the standpoint of formal logic and mathematics, such incompatibility appears exceedingly implausible, for the simple reason that in these fields human reasoners make conscious, explicit use of rules of syntactic/proof-theoretic inference, *and* diagrams that depict precisely the kind of visuo-spatial objects Johnson-Laird believes are at the very heart of human reasoning. Empirical evidence of such heterogeneous reasoning can be found on the pages of any number of journal articles and textbooks in these fields.

[4] Of course, a visual interface *per se* isn’t novel. After all, Toulmin’s scheme is diagrammatic. But Slate’s workspace includes not just a visual representation for inferential links between propositions, but also: visual depictions of models, relational databases, ontologies, and so on.

[5] *Rationale* (Austhink Software Pty Ltd., 2006–2007) is an argument diagramming software system.

[6] We use “children” in the typical graph-theoretic fashion, but atypical of directed graphs: arrows here point *from* children *to* parents.

[7] See `http://kryten.mm.rpi.edu/Vivid.pdf`.

[8] Phase 1 is currently performed by the user, but techniques are in development, and proposals under review for funding, to automate this phase

Phase 2: Discourse representation structures (DRSs) are automatically generated from the controlled English. DRSs are a syntactic variant of first-order logic for the resolution of unbounded anaphora. Their use in the interpretation of text is a central element of discourse representation theory [27, 28].

Phase 3: The DRSs are automatically translated [29] into MSL, the chief native language of Slate. As a DRS is equivalent to a quantified first-order formula, translations [3] to first-order logic and MSL are not conceptually difficult.

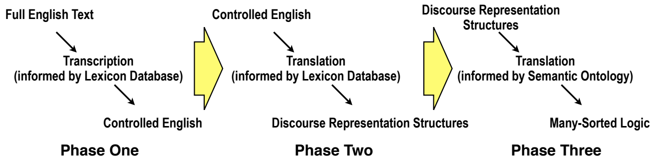


Figure 7. Slate's natural language understanding process.

Slate also presents arguments linguistically. Slate builds English justifications from arguments in the workspace, and from machine-generated proofs. Arguments and natural deduction proofs are already poised for efficient translation into English. They require no further document structuring or content determination.[9]

Slate's natural language generation system decomposes a proof or argument into constituent subproofs, translating each subproof from the top down. Once a subproof is translated, it is sent to a microplanning system that maps subproofs to discourse relations [23]. Though the overall structure of the proof or argument must remain the same, this restriction is not imposed on subproofs, which are molded and fitted to a number of different discourse relations and rhetorical structures for the sake of fluidity (e.g., redundancy checking, referring expression generation). In this fashion, Slate typically generates well-structured, expressive expositions.[10]

4.3 Reasoning Technology

Though human reasoners are superb at concurrently employing many different reasoning techniques (i.e., at heterogeneous reasoning), and are creative, they are also susceptible to bias and cognitive illusions. Human reasoners are also limited by the size of working memory, and so find it extremely difficult to keep large argument structures consciously in mind,

Machine reasoning systems, then, are a perfect complement to the human reasoner. Though a machine reasoning currently lacks creativity and is often limited to one mode of reasoning, computer memories and processors are vast and fast, making machine reasoning systems excel at the mechanical, brute-force reasoning tasks that their human counterparts find so difficult and tedious.

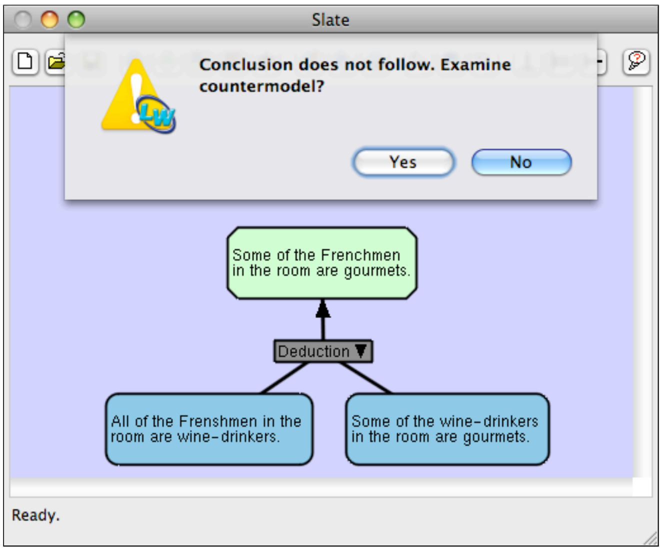
Slate harnesses these strengths by incorporating some of the best machine reasoning systems available, such as theorem provers and model finders. We describe three ways in which Slate exploits machine reasoning systems:

Argument checking: An inference in Slate is tagged with a rule or reason that justifies that inference. These individual inferences are checked for correctness using *Athena* [1], a type-Ω Denotational Proof Language (DPL) [2]. DPLs allow Slate to the application of rules with fine granularity. DPLs allow users of Slate to define new rules and reasons. This technique is useful in abbreviating common inference sequences, and for domain-specific reasoning.

Proof search: Slate uses two resolution based reasoners for proof search: SRI's SNARK, a multi-sorted reasoning kit implemented in Lisp; and OTTER [49], a single-sorted prover written in C. Pollock's OSCAR [37] is also used to find natural deduction proofs that are suited for human consumption.

Model finding: Slate finds models with PARADOX [16], a model finder that won the SAT/Models class in the 2003 CASC. Paradox finds finite models which are rewritten in AT&T's `dot` language which GraphViz renders as a visual model (e.g., Figure 3).

The benefits of these systems are great, yet most argument mapping systems seem to lack the features that machine reasoning systems provide. For instance, the argument mapped in *Rationale* in Figure 6 is a cognitive illusion.[11] In Slate, a counterexample would be generated that shows the possible state wherein all Frenchmen are wine drinkers, but the only gourmets in the room are non-French wine drinkers (Figure 8). For an algorithm able to decide whether or not any syllogism is valid, see [12].



Conclusion does not follow. Examine countermodel?

Some of the Frenchmen in the room are gourmets.

All of the Frenshmen in the room are wine-drinkers.

Some of the wine-drinkers in the room are gourmets.

Figure 8. Slate reveals the illusion in Figure 6's argument.

4.4 Uncertainty & Strength Factors

Human reasoners often speak in what may initially appear to be the language of probability theory. We say that one outcome is "more likely" than another, that some occurrence "isn't likely," that evidence is "fairly reliable," and that a conclusion is "beyond reasonable doubt." Yet psychological research confirms (e.g. see [26]) that human cognition is *not* based in probability theory. Slate reflects our belief that such phrases as those just quoted reflect *strength factor*-based reasoning, in which propositions and arguments are tagged with strengths from a discrete nine-point spectrum ranging from

---

as well [30, 5, 4]. Expertise presently needed in order to accomplish the rephrasing in question must include thorough knowledge of the syntax of ACE/CELT. Typically, the challenge is to preserve meaning while sacrificing sophisticated tense and model content.

[9] That is, document planning [38] is *prima facie* complete.

[10] Nonetheless, predictably, we have not yet reached the point at which the elegance of informal proofs typical of journals in the formal sciences is present in Slate-generated English.

[11] This particular illusion is known as the "wine drinker" illusion, introduced in [24]. Please note that abduction is not applicable in this problem, as the task is to determine whether the conclusion in question *must* be true if the premises are. As is well-known, to enforce the need for correct reasoning on the part of subjects faced with reasoning problems, one need only inquire as to whether a purported conclusion is logically necessitated by the premises in question.

'Certainly False' to 'Certain,' and that includes the terms 'Beyond Reasonable Doubt,' 'Probable,' and 'Counterbalanced.'

Strength factors are applied manually to individual inferences and to propositions that are not the conclusion of any inference. Strength factors are then propagated through the entire argument structure. This regulated use of strength factors in Slate helps users to pinpoint weak points in complex arguments, and allows them to focus their efforts effectively.

4.5 Interoperability

Slate's practical usefulness is greatly enhanced by virtue of interoperability with external data-stores (e.g., knowledge-bases, databases) and analytical tools. Slate achieves this through *provability-based semantic interoperability* (PBSI) [46, 43], a framework designed to facilitate information exchange between fundamentally different knowledge-management systems, provided only that there is some meaningful (i.e., semantic) relationship between the information in the systems' data-stores. Various standards and languages (e.g., XML) have been used for *syntactic* interoperability, but until PBSI there had been no general method for *semantic* interoperability. PBSI is designed to exchange not only information that is differently *structured* (e.g., XML, relational databases, text files), but knowledge that differs *semantically*.

PBSI includes a language for *bridging axioms* that formalize the relationships between ontologies. An extension, PBSI$^+$, associates with each information exchange a proof certifying the conservation of meaning. *Translation graphs* build bridges between ontologies that preserve semantics. A translation graph is a directed graph whose vertices are ontologies, and whose edges are atomic axiomatic relationships between ontologies (Figure 9).

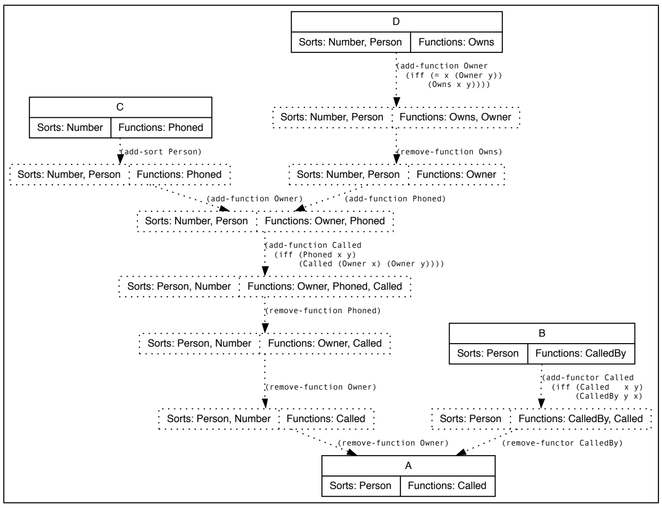


Figure 9. A translation graph generates the axioms facilitating semantic interoperability between four ontologies.

The framework and corresponding implementation of PBSI significantly reduces the work required for semantic interoperability, and partially automates the process.[12] PBSI has enabled interoperability between Slate and a number of commercial and academic analytical declarative representation schemes (e.g., SQL, Common Logic (now an ISO standard), and OWL).

5 EXAMPLES

The initial example (§2) showed some of the capabilities of Slate (e.g., model finding, inference validation), but glossed over the construction of the final argument (Figure 5). We now present two examples that highlight the *process* of building arguments in Slate. The first example is strictly formal, and motivated by well-known results that have not yet been formalized and mechanically validated; the second recreates an argument from Sherlock Holmes.[13] But before presenting the pair of examples, a note on the centrality of deductive reasoning.

5.1 The Centrality of Deductive Reasoning

Slate is designed to assist human reasoning in all its established modes (deductive, inductive, abductive, etc.). Nonetheless, in the present paper, space doesn't allow coverage of all these modes—and not only that, but deduction, it must be conceded, gets top billing herein. Why is that? There are a number of reasons; we give four.

First, while it's easy enough for humans to articulate exclusively deductive reasoning (e.g., in published proofs in logic, mathematics, computer science, etc.), it is well nigh impossible for humans to produce high-quality argumentation that is exclusively in a non-deductive mode, that is, argumentation without, at least in part, employing deduction. For example, opinion pieces in newspapers appear daily in the tens of thousands across the globe, and each and every one makes use of deduction; perhaps not extensive use, but use nonetheless. Second, while there is no consensus as to what valid (say) inductive reasoning is, there is universal agreement that certain patterns of deductive reasoning (e.g., *modus tollens*, proof by cases, etc.) provide a canonical standard for valid deductive reasoning. Third, since some entire disciplines are based on deduction (e.g., computer science, based on deductive logic; see e.g. [20]), and since we wish Slate to be useful to practitioners in these fields, deduction is important to us. Fourth and finally, it is natural that emphasis be placed upon deductive reasoning, for the simple reasoning that children in the civilized world are premeditatedly exposed to deduction (and expected to learn certain elementary forms of it) in K–12 mathematics education. By contrast, while in such education explicit coverage of deduction is often mandated (e.g., in the United States, some states, e.g., New York, require students to be explicitly taught deductive logic from first grade on), such is never the case, as far as we know, for induction, abduction, analogical reasoning, and so on.

5.2 Formal Reasoning

Gödel's first incompleteness theorem (GI), one of the most celebrated results in mathematical logic, states that any consistent, formal, recursively enumerable theory (e.g., a first-order logical system) that proves basic arithmetical truths cannot prove *all* arithmetical truths [21]. This result is often difficult to accept upon initial exposure, and we attribute this incredulity to: (i) the novelty and ingenuity of Gödel's proof, and (ii) the counter-intuitiveness of his result. Following in the pedagogical tradition of Smullyan [44], we present

---

[12] Currently, PBSI is semi-automated. Work is underway on a full automation, but the task is sobering, since such automation will require considerable progress in an area of AI and computer science that has been stalled since its inception: automatic programming. Bringsjord & Arkoudas currently have an exploratory grant from NSF to explore this direction.

[13] More demonstrations and examples are available from the Slate website, `http://www.cogsci.rpi.edu/slate/`, and demonstrations page, `http://www.cogsci.rpi.edu/slate/demos`.

a Gödelian logic puzzle that approximates GI and demonstrates the power of Slate within demanding logico-mathematical domains like those in which Gödel worked.

A Precursor Gödelian Puzzle. Suppose a machine $\mathcal{M}$ operates on expressions: finite, non-empty sequences of the four glyphs $\sim$, $\star$, P, and M. These four glyphs have intuitive meanings: $\sim$ stands for 'not,' $\star$ for 'to be' or 'is,' P for 'provable,' and M for 'mirror of,' where the mirror of an expression ϕ is the expression $\phi \star \phi$. A sentence is an expression of a particular form, also with an intuitive meaning, specifically,

> $P \star \phi$ means that ϕ is provable and is true if and only if ϕ is provable by $\mathcal{M}$.
>
> $PM \star \phi$ means that the mirror of ϕ is provable, and is true if and only if the mirror of ϕ is provable by $\mathcal{M}$.
>
> $\sim P \star \phi$ means that ϕ is not provable, and is true if and only if ϕ is not provable by $\mathcal{M}$.
>
> $\sim PM \star \phi$ means that the mirror of ϕ is not provable, and is true if and only if the mirror of ϕ is not provable by $\mathcal{M}$.

$\mathcal{M}$ is such that it only proves true sentences and never false sentences (i.e., the machine is *sound*). Prove that $\mathcal{M}$ cannot prove all true sentences—there is a true sentence which cannot be proved by $\mathcal{M}$ (i.e., the machine is *incomplete*).

Formalization of the Gödelian Puzzle. We formalize the above puzzle as a logical language consisting of the constants: $\sim$, $\star$, P, M; the (unary) predicates: glyph, expression, sentence, provable, and true; and the functions: cat (concatenation), and mirror. For convenience, we describe as glyphs, expressions, sentences, provable, and true any terms on which glyph, expression, sentence, provable, and true holds, respectively, and denote the application of cat to two terms ϕ and ψ as the concatenation of ϕ and ψ, or by $\phi\psi$, and the application of mirror to a term ϕ as the mirror of ϕ. The interpretation of this vocabulary is subject to the following twelve axioms:

1. *The constants $\sim$, $\star$, P, and M are each distinct.*
2. *The constants $\sim$, $\star$, P, and M are the only glyphs.*
3. *The concatenation of two terms is an expression if and only if both terms are themselves expressions.*
4. *Concatenation is associative.*
5. *The term ϕ is an expression if and only if ϕ is a glyph or is the concatenation of two expressions.*
6. *The mirror of an expression ϕ is defined as the concatenation of ϕ, $\star$, and ϕ (i.e., $\phi \star \phi$).*
7. *If ϕ is an expression, then $P \star \phi$, $PM \star \phi$, $\sim P \star \phi$, and $\sim PM \star \phi$ are sentences.*
8. *If ϕ is an expression then the sentence $P \star \phi$ is true if and only if ϕ is provable.*
9. *If ϕ is an expression, then the sentence $PM \star \phi$ is true if and only if the mirror of ϕ is provable.*
10. *If ϕ is an expression, then the sentence $\sim P \star \phi$ is true if and only if ϕ is not provable.*
11. *If ϕ is an expression, then the sentence $\sim PM \star \phi$ is true if and only if the mirror of ϕ is not provable.*
12. *Every sentence ϕ that is provable is also true.*

The given axioms (propositions 1–12) are represented visually in the Slate workspace in Figure 10), each consisting of the first-order formula derived from the English descriptions above. Moreover, a new intermediate hypothesis is introduced toward the desired goal, viz., that there is a true sentence that cannot be proved by $\mathcal{M}$:

13. *$\sim PM$ is an expression.*

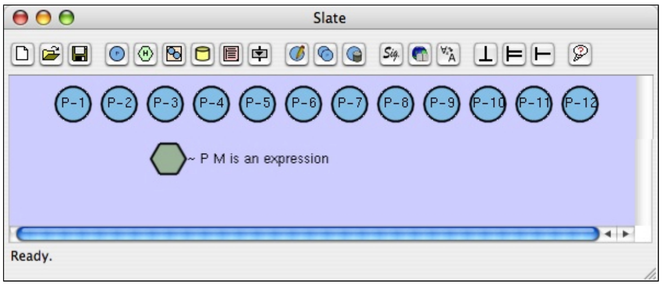


Figure 10. Propositions 1–12 and hypothesis 13 in the Slate workspace.

We indicate that hypothesis 13 is a logical consequence of propositions 2, 3 and 5 by drawing a deductive inference from each of these propositions to hypothesis 13 (Figure 11). Slate is then able to confirm or refute the added inference. Slate does indeed confirm that hypothesis 13 follows from the indicated propositions, by producing as evidence a formal proof which is added to the workspace as a *witness*. Witnesses are objects in Slate that support or weaken inferences. The double-plus symbol indicates that the witness confirms the argument, an ability reserved only for formal proofs. If the inference had been invalid, Slate might have produced a countermodel demonstrating the inference's invalidity.

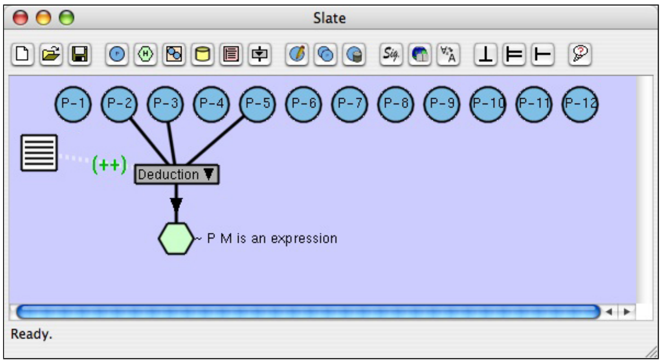


Figure 11. Proof of $\{2,3,5\} \vdash 13$ in the workspace and verified by Slate.

Having proved $\sim PM$ *is an expression*, it follows from 13 and 7 that:

14. *$\sim PM\star \sim PM$ is a sentence.*

If we suppose that $\sim PM\star \sim PM$ *is not true* then by 11 *the mirror of* $\sim PM$ *is provable* and thus by 6 $\sim PM\star \sim PM$ *is provable*. But then, according to 13 and 14, $\sim PM\star \sim PM$ *is true*—which is in contradiction with our supposition that $\sim PM\star \sim PM$ *is not true*. And so it must be the case that $\sim PM\star \sim PM$ *is true*. In other words, as shown in Figure 12, the hypothesis that

15. *$\sim PM\star \sim PM$ is true.*

follows from axioms 6 and 11 and hypotheses 12 and 13. Since $\sim PM\star \sim PM$ *is true*, it follows from 6 and 11 that

16. *$\sim PM\star \sim PM$ is not provable.*

and consequently, that *there is a true sentence which cannot be proved* (Figure 13).

5.3 Informal Reasoning

When using Slate, the reasoner is able to construct arguments that more closely resemble the uncertain and informal nature of everyday, natural inference. Moreover, the user benefits from the system's

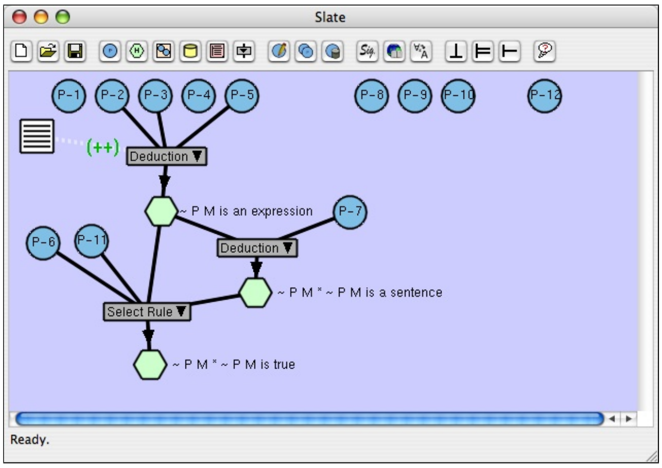


Figure 12. Argument that $\sim PM\star\sim PM$ *is true*.

automated evaluation and strength propagation mechanisms. To illustrate these capabilities, we recreate an argument given by Sherlock Holmes in the Sir Arthur Conan Doyle's short story *The Adventure of the Blue Carbuncle*, wherein the famous detective discovers some qualities of the owner of a lost hat. Upon inspection of the hat, Holmes and Watson observe that it is lined with red silk, has a flat brim with curled edges, and a hat-securer (for protecting against the wind) that has fallen into disrepair (Figure 14).

Holmes' observations of the hat are the result of inspection and, therefore, for our purposes, are certain. Yet Holmes continues and *abduces*, (i) from the observation that the hat has a flat brim with curled edges, and (ii) his knowledge that such hats were in style three years prior, that (iii) the hat is three years old.[14] We add Holmes' background knowledge to the Slate workspace and construct Holmes' abductive inference (Figure 15), marking the inference as Abductive (abbreviated as Abd), and we assign it the strength (1) Probable. Probable is the weakest positive strength factor, indicating only that the inference is more likely to be correct than incorrect. The strength of the inference propagates toward the hypothesis. Each of the premises are certain, and the inference is Probable, so the conclusion receives the strength Probable.[15]

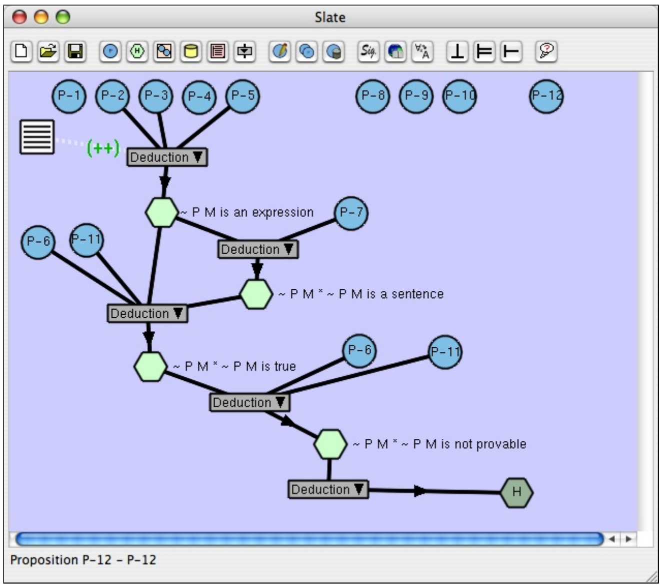


Figure 13. Argument that *there is a true sentence which cannot be proved*.

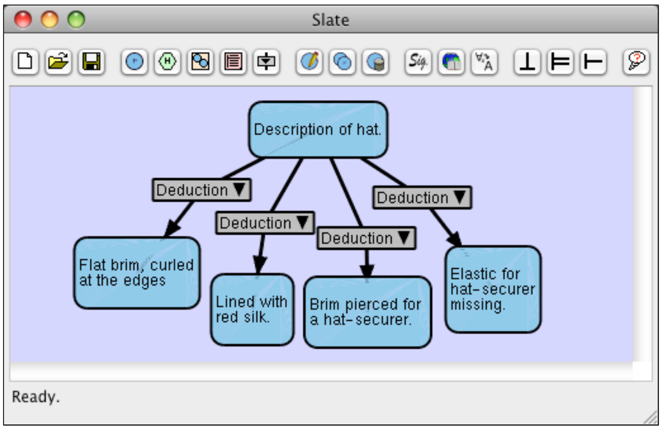


Figure 14. Observed properties of the found hat.

Holmes proceeds to infer that, although the owner of the hat was well off at the time the hat was purchased, the fact that the owner has worn the same (now tattered) hat for three years suggests that he has suffered financial losses (Figure 15). Holmes would admit that this conclusion is not an absolute certainty; several alternate hypotheses could explain why the owner wore this particular hat, e.g., that this hat held sentimental value, or that the owner rushed to leave the house and this was the first hat he came upon. Alternate explanations are commonplace in abductive reasoning and their plausibility is the basis for ascribing strengths to abductive inferences.

That the owner had a hat-securer installed indicates, to Holmes at least, that the gentleman possessed a high degree of foresight; but that he failed to repair the elastic when it broke suggests he has less foresight now than he did in years past. This, along with the hypothesis that the owner has suffered financial difficulties, leads Holmes' to conclude that the owner has fallen into moral decline (Figure 16).

6 SLATE'S EFFICACY

Setting aside the theoretical issues, one might ask if Slate effectively assists humans in reasoning formally. To gauge this, an experiment was conducted in which humans were required to use either the Slate software or pencil and paper to solve reasoning problems.[16] Results

---

[14] We haven't space available to explain abduction in Slate, and note only, first, that abduction, whether in or or out of Slate, conforms, in general, to the (deductively invalid) schema that from $\phi \rightarrow \psi$ and ψ one infers to ϕ. Second, abduction in Slate is based more specifically upon the following core concept: Let Φ be some set of formulas, and ψ some individual formula. Suppose that there is an algorithm $\mathcal{A}$ able to ascertain whether or not ψ can be proved from Φ (i.e., whether or not $\Phi \vdash \psi$). Then automated abductive reasoning consists in the running of a composite algorithm $\mathcal{A}^\star$ that first runs $\mathcal{A}$ on a relevant pair (Φ, ψ), and if a negative verdict is returned, searches for a formula α such that $\Phi \cup \{\alpha\} \vdash \psi$. If such a formula can be obtained, then it is (abductively) inferred. The reader should satisfy herself that Sherlock Holmes can be plausibly viewed as running the $\mathcal{A}^\star$ algorithm. This is as good a place as any to point out that all abductive (and, for that matter, inductive) reasoning can be recast as enthymematic deduction.

[15] The propagation algorithms in Slate, frankly, are not straightforward, and are beyond the relatively small amount of space we have available here.

[16] None of these problems are of a type seen repeatedly in everyday reasoning, outside the formal sciences, or training therein. However, Slate has

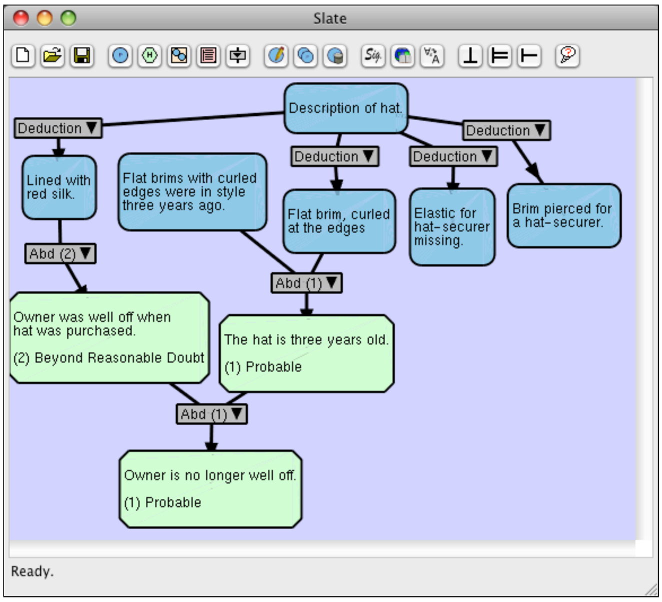


Figure 15. Some intermediate conclusions drawn (abductively) from the given facts about the hat.

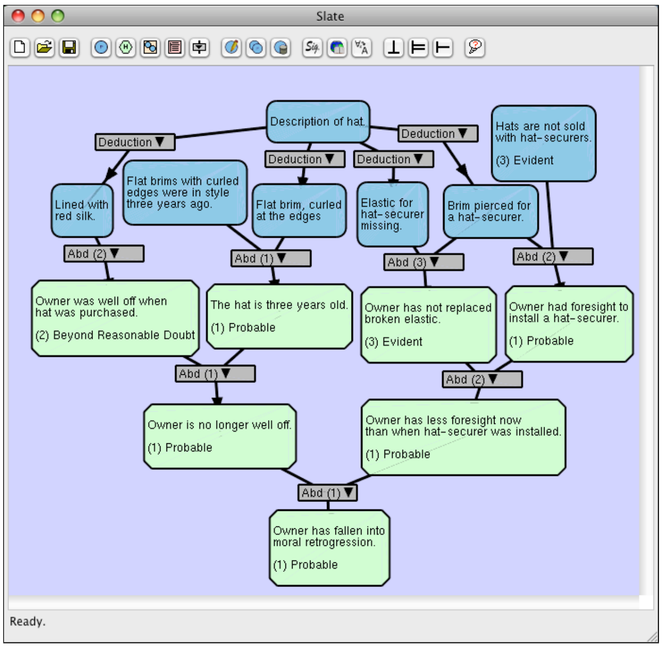


Figure 16. An abductive argument contending that the owner of the missing hat has fallen into 'moral retrogression.'

demonstrate that humans perform far better with Slate than they do on their own. We briefly describe the experiment herein.

Methodology. Forty undergraduates from Rensselaer Polytechnic Institute served as participants. They formed two groups for each condition of the experiment, with twenty subjects in each group. The experiment was run on paper, as well as on two Apple Power Mac G5s with 20" flat screen monitors.

Design. The experiment was a univariate between-subject design where subjects were told to complete a number of reasoning problems. The independent variable had two conditions: subjects either solved the reasoning problems by hand using pencil and paper (the control condition), or were allowed to use Slate to help solve the problems. The salient dependent variable, and the metric for good performance, was the participant's solution accuracy.

Procedure. Subjects were given a series of reasoning problems.[17] Each problem was of the following format. A number of premises $a_1, a_2, \ldots, a_n$ would be enumerated, followed by a prompt p (viz., "What logically follows?"). Subjects were told to provide an answer to p and a corresponding justification for their answer. They were allowed as much time as they needed to finish the problems, but were not allowed any outside help or reference. Subjects in the paper and pencil condition were talked through an example inference problem and then allowed to begin their work on the problems in the experiment. Subjects using Slate were shown a video of the system that demonstrated various relevant features of the software as a sample problem was solved (subjects had no prior exposure to Slate). Once subjects felt comfortable with their task, they were allowed to begin. Subjects were tested on three problems of varying difficulty.

Results. For brevity, we will only look at the most cognitively difficult problem given to the subjects, one of Johnson-Laird's *king-ace* "illusory inference" problems [24, 11] (Figure 17).

Exactly one of the following statements is true:

- If there is a king in hand, then there is an ace in hand.
- If there isn't a king in hand, then there is an ace in hand.

What can you infer from the above premise?
Provide an answer and a justification:

Figure 17. The *king-ace* problem (as presented to subjects).

The problem is notoriously difficult, as the solution defies (logically untrained) intuition. That is, naïve intuition leads to the answer that there is an ace in the hand, though it follows deductively that there *cannot* be an ace in the hand.[18] Individuals overwhelmingly

---

been used with great success by humans tackling robust case studies in the realm of intelligence analysis.

[17] We haven't the space to go through the problems used. In addition to the one upon which we momentarily focus, we mention that the problem of ascertaining what can be deduced from 'It's not the case that: If the cat is black, it sits upon the mat.' was investigated, and that the profound difficulty of this problem for even logically untrained but graduate-degree-holding people is at the heart of a famous mental logic system (PSYCOP) intended to model naïve human reasoning [40]. The answer is: 'That the cat is black.'—and for that matter that 'The cat doesn't sit upon the mat.' Armed with Slate, most subjects weren't fooled.

[18] To see why there cannot be an ace, recognize that one of the two conditionals in the premise must be false, and a conditional is false only when its antecedent is true and its consequent is false. If "if there is a king in the hand, then there is an ace" is false, then there is not an ace, and if "if there isn't a king in the hand, then there is an ace" is false, then again, there is not an ace. So regardless of which conditional is false, there cannot be an ace in the hand. QED Of course, this proof notwithstanding, there's no denying that 'There is an ace in the hand' *looks* to be deducible. Johnson-Laird's explanation for this (of course) is rooted in mental models, but a simpler explanation is simply that subjects perceive proof by cases to be appropriate in this case. Finally, logically untrained subjects are notoriously bad at deducing the truth of the antecedent and the falsity of the consequent from a negated conditional, and there is no shortage of psychological explanations as to why, from quarters other than the mental models camp.

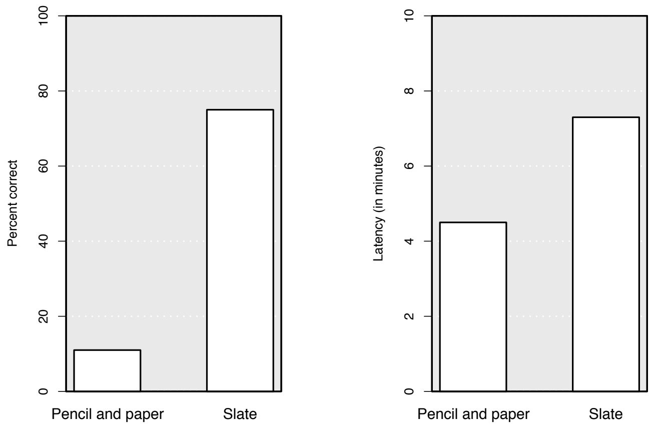


Figure 18. Average times and scores for the *king-ace* problem. While subject using Slate took more time to complete the task, their scores were markedly higher.

succumb to their System 1 intuitions instead of coming to the correct solution, and confirming it with a proof. Can Slate assist reasoners in solving this problem correctly?

Subjects using Slate were much more likely to correctly answer the king-ace problem. Using a single-factor ANOVA, we found that there was a significant effect in the performance on the aforementioned problem when assisted by Slate (M=75%, SE=1.25) than when not (M=11%, SE=3.86), F(1)=25.18, $p < .005$. Seventy-five percent of subjects solved the problem correctly using Slate, while only eleven percent managed to solve the problem without Slate's help (Figure 18). In addition, there was a significant effect in the performance time between the two conditions on the king-ace problem. Using another single-factor ANOVA, we found that subjects using Slate took longer to complete the problem (M=7.31, SE=68.86) than those using pencil and paper (M=4.5, SE=68.86), F(1)=9.39, $p < .005$.

Discussion. The data shows that Slate helps users avoid failing prey to "illusory inferences" (at least in the king-ace problem). As for the response time differences, we note that upon reading the problem, students quickly fall into the trap. Using only pencil and paper, they do little but elaborate on their (flawed) initial reasoning. Using Slate, they are presented with countermodels and are exposed to 'correct' reasoning. Processing these models, of course, takes some time.

This initial test produced encouraging results for Slate, but more experiments are needed to definitively evaluate Slate's efficacy at assisting users (after all, it is difficult to make broad statements about the system's efficacy from one experiment). Further experimentation is planned, and we anticipate that the results will demonstrate that users of Slate easily overcome a number of well-known biases in human reasoning and decision-making [35, 36].

A final pair of discussion points: We anticipate that some readers will find the king-ace problem deficient from either or both a pragmatic and/or educational point of view. From the former perspective, a reader might proclaim that in the real world of natural argumentation, no one would bother to tell someone two propositions if exactly one of them is known to be false. Actually, this isn't true. Everyday reasoning is rife with examples in which a pair of propositions is exclusively disjoined, and some conclusion is supposed to follow therefrom. For example, we might be informed by Jones that either a country from Europe will win the World Cup, or one from South America will (but not both, obviously). Jones might then proceed to argue that no matter which case obtains, the goalie on the victorious team will not have secured glory for an Asian country. From an educational point of view, a reader might express concern that Slate users who solved king-ace didn't need to *explain* the solution. In response, our objective is first to give humans a machine partner able to surmount difficult logical illusions (and the like). Later, we do intend to investigate whether Slate can function as efficacious teaching software.

7 FUTURE WORK

The development of Slate is ongoing. Below are a few of the enhancements that are currently in the works.

- Slate's implementation is being refactored. The next version will feature a new Java-based UI; separation of UI client from automated reasoning back-end so that the latter can be located on high-performance servers; and a 'shared workspace' so that multiple users can collaborate on analysis and arguments.
- In addition to support for shared workspaces, we are exploring ways to exploit Semantic Web technologies including RDF datastores and OWL ontologies. RDF might be used as an interchange format for Slate models: Slate could exchange models and hypothesis with other Semantic Web systems.
- As already noted, Slates makes use of various model-finding technologies. Unfortunately, extant model-finding systems are limited to finite domains. Work is underway to expand the reach of Slate's model-finding capabilities to certain well-behaved infinite domains [42].
- Slate will soon include support for so-called "visual logics," such as Arkoudas & Bringsjord's VIVID.[19] Visual logics combine diagrammatic and symbolic reasoning for any computable image.

ACKNOWLEDGEMENTS

We are greatly indebted to three anonymous referees for insightful objections and suggestions, and to Sangeet Khemlani for psychology contributions. With much gratitude, we acknowledge the financial support provided by ARDA's Novel Intelligence from Massive Data (NIMD) program, and also to ARDA's successors: DTO and (the agency operative today) IARPA, for continued support under the AQUAINT and ASpace-X programs. We would thank the many brilliant and dedicated intelligence analysts who continue to work with us. Finally, we are indebted to numerous colleagues at RPI for their insights and assistance.

REFERENCES

[1] K. Arkoudas, 'Athena'. http://www.cag.csail.mit.edu/˜kostas/dpls/athena.

[2] K. Arkoudas, *Denotational Proof Languages*, Ph.D. dissertation, MIT, Department of Computer Science, Cambridge, USA, 2000.

[3] P. Blackburn and J. Bos, 'Working with Discourse Representation Theory'. Forthcoming.

[4] J. Bos, 'Towards wide-coverage semantic interpretation', in *Proceedings of the 6th International Workshop on Computational Semantics*, pp. 42–53, Tilburg, The Netherlands, (2005).

[5] J. Bos, S. Clark, M. Steedman, J. R. Curran, and J. Hockenmaier, 'Wide-Coverage Semantic Representations from a CCG Parser', in *Proceedings of the 20th International Conference on Computational Linguistics*, Geneva, Switzerland, (August 23–27 2004).

[6] *Mental Logic*, eds., M. D. S. Braine and D. P. O'Brien, Lawrence Erlbaum, Mahwah, NJ, 1998.

---

[19] Specification of VIVID is available at: `http://kryten.mm.rpi.edu/Vivid.pdf`.

[7] S. Bringsjord, K. Arkoudas, M. Clark, A. Shilliday, J. Taylor, B. Schimanski, and Y. Yang, 'Reporting On Some Logic-Based Machine Reading Research', in *Machine Reading: Papers from the AAAI Spring Symposium Technical Report* SS-07-06, ed., O. Etzioni, pp. 23–28. AAAI Press, (2007).

[8] S. Bringsjord, E. Bringsjord, and R. Noel, 'In Defense of Logical Minds', in *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, eds., M. A. Gernsbacher and S. J. Derry, pp. 173–178, Mahwah, NJ, (1998). Lawrence Erlbaum.

[9] S. Bringsjord and D. Ferrucci, 'Logic and artificial intelligence: Divorced, still married, separated...?', *Minds and Machines*, **8**, 273–308, (1998).

[10] S. Bringsjord and D. Ferrucci, 'Reply to Thayse and Glymour on Logic and Artificial Intelligence', *Minds and Machines*, **8**, 313–315, (1998).

[11] S. Bringsjord and Y. Yang, 'Cognitive Illusions and the Welcome Psychologism of Logicist Artificial Intelligence', in *Philosophy, Psychology, and Psychologism: Critical and Historical Readings on the Psychological Turn in Philosophy*, ed., D. Jacquette, 289–312, Kluwer Academic Publishers, Dordrecht, The Netherlands, (2003).

[12] S. Bringsjord and Y. Yang, 'Problems used in psychology of reasoning are too easy, given what our economy demands', *Behavioral and Brain Sciences*, **26**(4), 528–530, (2003).

[13] Selmer Bringsjord, 'Declarative/Logic-Based Cognitive Modeling', in *The Handbook of Computational Psychology*, ed., Ron Sun, 127–169, Cambridge University Press, Cambridge, UK, (2008).

[14] Selmer Bringsjord, Deepa Mukherjee, Gerwin Schalk, Peter Brunner, and Sean Austin, 'EEG-evinced neuroanatomical correlates of the distinction between formal, domain-independent and concrete, domain-dependent deductive reasoning', (forthcoming).

[15] R. M. Chisholm, *Theory of Knowledge*, Prentice Hall, Englewood Cliffs, NJ, 1966.

[16] K. Claessen and N. Sorensson, 'New techniques that improve Mace-style finite model finding', in *Model Computation: Principles, Algorithms, Applications (Cade-19 Workshop)*, Miami, FL, (2003).

[17] N. E. Fuchs, U. Schwertel, and R. Schwitter, 'Attempto Controlled English (ACE) Language Manual, Version 3.0', Technical Report 99.03, Department of Computer Science, University of Zurich, Zurich, Switzerland, (1999).

[18] M. Genesereth and N. Nilsson, *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, Los Altos, CA, 1987.

[19] Vinod Goel, 'Chapter 20: Cognitive Neuroscience of Deductive Reasoning', in *The Cambridge Handbook of Thinking and Reasoning*, eds., Keith Holyoak and Robert Morrison, 475–492, Cambridge University Press, Cambridge, UK, (2005).

[20] Joseph Halpern, Robert Harper, Neil Immerman, Phokion Kolaitis, Moshe Vardi, and Victor Vianu, 'On the unusual effectiveness of logic in computer science', *The Bulletin of Symbolic Logic*, **7**(2), 213–236, (2001).

[21] M. Hirzel. On formally undecidable propositions of *Principia Mathematica* and related systems I, November 2000. Accessed electronically, http://www.research.ibm.com/people/h/hirzel/papers/canon00-goedel.pdf, Jan., 2007.

[22] S. Hoefler, 'The Syntax of Attempto Controlled English: An Abstract Grammar for ACE 4.0', Technical Report ifi-2004.03, Department of Informatics, University of Zurich, Zurich, Switzerland, (2004).

[23] E. H. Hovy, 'Automated discourse generation using discourse structure relations', *Artificial Intelligence*, **63**(1–2), 341–385, (1993).

[24] P. N. Johnson-Laird, 'Rules and Illusions: A Critical Study of Rips's *The Psychology of Proof*', *Minds and Machines*, **7**(3), 387–407, (1997).

[25] P. N. Johnson-Laird, *How We Reason*, Oxford University Press, New York, NY, 2006.

[26] *Choices, Values, and Frames*, eds., D. Kahneman and A. Tversky, Cambridge University Press, Cambridge, UK, 2000.

[27] H. Kamp and U. Reyle, *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Springer, 1993.

[28] H. Kamp and U. Reyle, 'A calculus for first order Discourse Representation Structures', *Journal of Logic, Language and Information*, **5**(3–4), 297–348, (1996).

[29] M. Manzano, *Extensions of First Order Logic*, Cambridge University Press, Cambridge, UK, 1996.

[30] D. Mollá and R. Schwitter, 'From Plain English to Controlled English', in *Proceedings of the 2001 Australasian Natural Language Processing Workshop*, eds., R. Schwitter, D. Estival, C. Paris, and A. Knott, pp. 77–83, Macquarie University, Sydney, Australia, (2001).

[31] N. Nilsson, 'Logic and Artificial Intelligence', *Artificial Intelligence*, **47**, 31–56, (1991).

[32] A. Pease and C. Fellbaum, 'An english to logic translator for ontology-based knowledge representation languages', in *Proceedings of the 2003 International Conference on Natural Language Processing and Knowledge Engineering*, pp. 777–783. IEEE Computer Society, (2003).

[33] A. Pease and C. Fellbaum, 'Language to logic translation with phrasebank', in *Proceedings of the Second International WordNet Conference*, eds., P. Sojka, K. Pala, P. Smrz, C. Fellbaum, and P. Vossen, pp. 187–192, Masaryk University, Brno, Czech Republic, (2004).

[34] F. J. Pelletier, 'Seventy Five Problems for Testing Automatic Theorem Provers', *Bell System Technical Journal*, **2**, 191–216, (1986).

[35] M. Piattelli-Palmarini, *Inevitable Illusions*, John Wiley & Sons, 1994.

[36] *Cognitive Illusions*, ed., R. Pohl, Psychology Press, New York, NY, 2004.

[37] J. L. Pollock, *Cognitive Carpentry*, MIT Press, Cambridge, MA, 1995.

[38] E. Reiter and R. Dale, *Building natural language generation systems*, Cambridge University Press, New York, NY, 2000.

[39] K. J. Rinella, S. Bringsjord, and Y. Yang, 'Efficacious Logic Instruction: People are not Irremediably Poor Deductive Reasoners', in *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, eds., J. D. Moore and K. Stenning, pp. 851–856, Mahwah, NJ, (2001). Lawrence Erlbaum.

[40] L. J. Rips, *The Psychology of Proof*, MIT Press, Cambridge, MA, 1994.

[41] *The Mental Models Theory of Reasoning*, eds., W. Schaeken, A. Vandierendonck, W. Schroyens, and G. d'Ydewalle, Lawrence Erlbaum, Mahwah, NJ, 2006.

[42] A. Shilliday, *Elisa : A New System for AI – Assisted Scientific Discovery Incorporating Novel Techniques in Infinite Model Finding*, Ph.D. dissertation, Rensselaer Polytechnic Institute, Troy, NY, forthcoming;2008.

[43] A. Shilliday, J. Taylor, and S. Bringsjord, 'Toward Automated Provability-Based Semantic Interoperability Between Ontologies for the Intelligence Community', in *Ontology for the Intelligence Community: Towards Effective Exploitation and Integration of Intelligence Resources (OIC-2007)*, ed., K. S. Hornsby, pp. 67–72, Buffalo, NY, (2007). National Center for Ontological Research.

[44] R. Smullyan, *Gödel's Incompleteness Theorems*, Oxford University Press, Oxford, UK, 1992.

[45] K. E. Stanovich and R. F. West, 'Individual Differences in Reasoning: Implications for the Rationality Debate', *Behavioral and Brain Sciences*, **23**(5), 645–665, (2000).

[46] J. Taylor, A. Shilliday, and S. Bringsjord, 'Provability-Based Semantic Interoperability via Translation Graphs', in *Advances in Conceptual Modeling - Foundations and Applications, ER 2007 Workshops CMLSA, FP-UML, ONISW, QoIS, RIGiM, SeCoGIS, Auckland, New Zealand, November 5-9, 2007, Proceedings*, volume 4802 of *Lecture Notes in Computer Science*, 180–189, Springer, (2007).

[47] S. E. Toulmin, *The Uses of Argument*, Cambridge University Press, Cambridge, UK, 1958.

[48] D. Walton, *Informal Logic: A Handbook for Critical Argument*, Cambridge University Press, Cambridge, UK, 1989.

[49] L. Wos, R. Overbeek, E. Lusk, and J. Boyle, *Automated Reasoning: Introduction and Applications*, McGraw Hill, New York, NY, 1992.

[50] Y. Yang and S. Bringsjord, 'Mental Metalogic: A New Paradigm for Psychology of Reasoning', in *Proceedings of the Third International Conference of Cognitive Science*, eds., L. Chen and Y. Zhuo, pp. 199–204, Hefei, China, (2001). USTC Press.

[51] Y. Yang and S. Bringsjord, 'Mental metalogic and its empirical justifications', in *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, eds., R. Alterman and D. Kirsh, pp. 1275–1280, Boston, MA, (2003). Cognitive Science Society.

[52] Y. Yang and S. Bringsjord, *Mental Metalogic: A New, Unifying Theory of Human and Machine Reasoning*, Lawrence Erlbaum, Mahway, NJ, forthcoming.

[53] Y. Yang, Y. Zhao, J. Zeng, J. Guo, Z. Ju, and S. Bringsjord, 'Empirical Justifications for the Universalness of the Mental Logic and Mental Models Paradigm', in *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, eds., B. G. Bara, L. Barsalou, and M. Bucciarelli, pp. 2399–2403, Mahwah, NJ, (2005). Lawrence Erlbaum.