Providing Decision Support for Cosmogenic Isotope Dating

Laura Rassbach, Elizabeth Bradley, and Ken Anderson
University of Colorado
Department of Computer Science
laura.rassbach@colorado.edu, lizb@cs.colorado.edu, kena@cs.colorado.edu

Abstract

Human experts in scientific fields routinely work with evidence that is noisy and untrustworthy, heuristics that are unproven, and possible conclusions that are contradictory. We present a fully implemented AI system, Calvin, for cosmogenic isotope dating, a domain that is fraught with these difficult issues. Calvin solves these problems using an argumentation framework and a system of confidence that uses two-dimensional vectors to express the quality of heuristics and the applicability of evidence. The arguments it produces are strikingly similar to published expert arguments. Calvin is in daily use by isotope dating experts.

Introduction

Automating scientific reasoning is an important challenge to AI. An automated tool can do boring and repetitive reasoning, freeing experts to do more difficult and creative work. Indirectly, it can make explicit the knowledge and reasoning used by experts in the field. Finally, an automated tool can consider all possibilities, sometimes exploring scenarios that human experts may miss.

This paper discusses automating reasoning for dating geological landforms. Dating landforms is similar to investigating a crime scene: from the information available on the surface, left behind by an unknown series of events, experts must abduce what happened in the past. In the example diagrammed in Figure 1, subsurface rocks are exposed over time as the soil around them erodes. A geoscientist would be faced with the situation shown on the right of the figure; his[1] goal is to derive the situation shown at the left, along with the processes that were at work and the timeline involved.

To accomplish this, a geoscientist first dates a set of rock samples from the present surface, then reasons backward to deduce what process affected the original landform. This is a difficult deduction: geological processes take place over an extremely long period of time, and evidence remaining today is scarce and noisy. Finally, experts in geological dating, like experts in any field, are only human, and can be biased in favor of one theory over another.

In the face of these problems, experts form an exhaustive list of possible hypotheses and consider the evidence for and against each one—much like the concept of argumentation. Our system to automate this reasoning, Calvin, uses the same argumentation process as experts, comparing the strength of the evidence for and against a set of hypotheses before coming to a conclusion. We collected knowledge about how isotope dating experts reason via interviews with several dozen geoscientists. Confidence is key in this kind of reasoning: not only in the quality of evidence, but also in the knowledge that is used to connect evidence to conclusion. Capturing these elements required a novel instantiation of confidence-based reasoning in an argumentation system. From these elements, Calvin produces arguments almost identical to the reasoning presented by human experts.

Calvin provides several contributions to AI and to the larger scientific community:

- Its rule base is an explicit representation of the knowledge of two dozen experts in landform dating
- It incorporates a rich system of confidence that captures the reasoning of real scientists in a useful way
- It is a fully implemented system in the beginning stages of deployment
- It is a real tool that is in use by real scientists

In the following section, we discuss the general problem of cosmogenic isotope dating, highlighting its challenges and the approach that experts take to solving it. Next, we describe how Calvin uses argumentation to automate that process, and finally, we discuss our results.

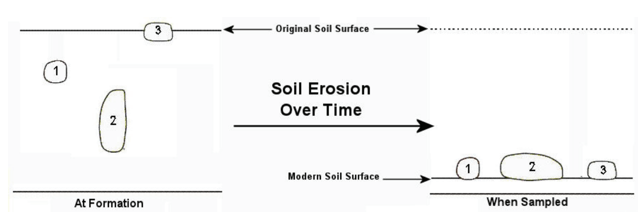

Figure 1: Deducing past events from the evidence available now.

[1]We use the male pronoun for linguistic simplicity.

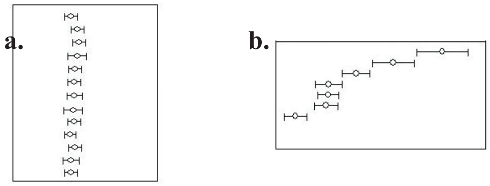

Figure 2: Possible sample age-error distributions
(a) An 'ideal' age distribution: many samples with a range of ages included in every single sample's age plus or minus the uncertainty in that sample's age.
(b) A 'real' age distribution: significant spread, fewer samples. This case is common in real problems.

Cosmogenic Isotope Dating

Beginning from a set of samples collected from boulders on a landform, an isotope dating expert's goal is to determine the absolute age of that landform. This section summarizes how experts work, from sampling individual boulders to deducing an age for an entire landform.

The first step is to collect as many samples as possible from the landform. A set of at least five samples is best (Putkonen and Swanson 2003); five to ten samples is about the norm. Experts would prefer to collect far more samples, but often only a handful of boulders suitable for sampling are available. While collecting samples, the expert also makes qualitative field observations that are often crucial for interpreting initial dating results.

Once the expert has gathered a set of samples in the field, he brings them to a lab for dating. He finds the exposure age of each sample by determining its isotopic composition (some isotopes are produced only by cosmic rays). Then he performs a series of calculations using this composition and some of the observations taken at the sample site (such as topographical shielding) to find the length of time the sample has been at the surface. This length of time takes the form of a value with error bars. The expert's next step is to derive an absolute landform age.

For most landforms, the surface exposure times of boulders on the surface are true measure of the age of the landform. This is because the boulders are brought to the surface from deep bedrock when the landform is formed. However, different landforms are exposed to different events, complicating the task of determining an overall age for a landform. The simplest version of this problem arises when the expert has a large number of samples and all of their ages overlap, as shown in Figure 2(a).

Unfortunately, sample sets rarely have a perfect range of overlap. Instead, initial sample ages are usually spread over a wider range than the individual sample errors as in Figure 2(b). In these cases, the researcher must construct an explanation for the spread in apparent ages, usually a geologic process acting on the samples over time. Once he has found a process that explains the majority of the data, he uses further calculations and educated guesswork to remove its effects from the sample set and, he hopes, arrive at a single age for the landform. In real landforms, more than one process may have been at work, but experts generally focus on isolating the one that most affected the ages of the samples.

Unfortunately, a single round of analysis does not always serve to isolate a landform's true age with any confidence. In this case, the expert must return to the original sample site (at great expense) to seek further samples that disambiguate between possible hypotheses or reinforce the evidence for a likely process. For example, a soil sample at depth can distinguish between several candidate processes.

Most explanations that experts use for a spread in apparent ages come from a short list of geologic processes that affect the exposure times of the samples. Statistical 'processes' may also explain the data: e.g. the age spread may be a result of lab error or some form of mis-sampling.

Despite the small number of candidate processes, selecting an explanation for the apparent age spread of a particular landform is not a simple task. Available data are noisy and untrustworthy. Experts make mistakes in their observations in the field. Moreover, the manifestation of one process may be quite similar to the manifestations of other processes. Experts make a final decision about the process in effect on the basis of heuristic reasoning. These heuristics frequently contradict each other, and different experts hold contradictory opinions about the correct heuristics. Addressing this contradiction was a major factor in Calvin's design, as we discuss in the following section.

Design and Architecture

Calvin's input is a set of samples that have already been individually dated (experts use a different tool for this step, such as ACE (Anderson et al. 2007)). It analyzes these groups of samples to determine what process affected the whole landform. Our selection of an appropriate framework for solving this problem rested primarily on data we gathered during extensive interviews. From these interviews, we arrived at an argumentation framework as the best one for Calvin. This selection led to a need to represent both expert knowledge and expert sources and comparison of confidence.

Interviews: Design Motivation

We interviewed two dozen experts in isotope dating, amassing around eighty hours of formal and informal interviews. Transcripts of formal interviews can be found in (Rassbach 2009). We learned that experts in isotope dating consistently use the method of multiple simultaneous hypotheses (Chamberlain 1965). They form arguments for and against every hypothesis, judging their relative and absolute strength to arrive at a solution.

We learned two significant things in our interviews. First, experts reason with contradictory heuristics:

> **Geologist:** The thing about inheritance is, it's usually thought about as quantized, not incremental...
> **Interviewer:** So it shouldn't be a spread of ages
> **Geologist:** Yeah, however, you can convince me you would see a continuum.

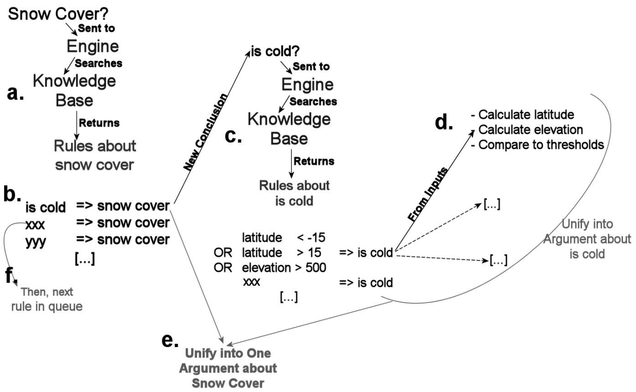

Figure 3: Illustration of Calvin's chaining of rules.
(a) Calvin's engine finds all the rules in its knowledge base with a 'snow cover' conclusion and puts them in an unordered set to consider one at a time.
(b) Calvin considers a rule that snow cover is more likely in cold areas. To apply this rule, Calvin must determine if the sampled area is cold, data not input directly.
(c) The main reasoning loop is called with a new conclusion, 'the area is cold.'
(d) Calvin sequentially considers every rule about coldness. We show a rule about coldness that finds the average latitude and maximum elevation of the sample site and compares those values to fixed thresholds.
(e) The results of arguing about 'is cold' are unified with the original rule about 'snow cover.'
(f) Calvin moves on to the next rule about snow cover.

That is, not only do experts disagree with each other, they sometimes disagree with *themselves*.

And second, experts themselves are convinced that reasoning in their field takes place in the form of argument:

Interviewer: So we're trying to understand what it is that you do

Geologist: Well, mostly we argue with each other.

The structure of expert reasoning revealed in these interviews makes argumentation a natural framework for automating expert reasoning.

Reasoning Process

Most processes that affect a landform come from a set list:

- The possibility that no process at all was at work
- Exhumation
- Clast erosion
- Inheritance
- Vegetation cover
- Snow cover
- The possibility that some sample(s) are outliers

Other processes do sometimes affect landforms, but these seven are the most common. Because the possible processes are known, experts do not generally need to form novel hypotheses to find an explanation for their data. Therefore, Calvin gives every hypothesis from the list of 'usual suspects' equal consideration, as recommended in (Chamberlain 1965) and by experts during our interviews.

Calvin's main task is generating arguments for and against each hypothesis in its list. This process involves finding the applicable information in its knowledge base, unifying it with sample data, and constructing a collection of arguments about the conclusion. Performing these functions requires a number of design elements: an engine, rules, evidence, and arguments.

Calvin considers candidate hypotheses one at a time and builds arguments for and against each hypothesis from the top down using backwards chaining. First, the engine finds all the rules that apply to this hypothesis—i.e., those that refer to the same conclusion. Unification is applied to each of these rules, resulting in either a new conclusion to consider or a comparison to input data. Calvin builds the most complete possible set of arguments from its knowledge base for and against each hypothesis.

Figure 3 illustrates this backwards-chaining process for an argument about the possibility of snow cover on a landform. Calvin's engine finds the applicable set of rules, considers each one in turn, and then forms a confidence in the overall evidence. Eventually Calvin will consider every rule about snow cover in its knowledge base and, if the data for unification exists, the rule will be used in its resulting reasoning.

Every rule in Calvin contains both a conclusion and a template for evidence that will support that conclusion. The primary portion of a rule is an implication of the form **A**=>**C**, where **A** may be either a single literal or the conjunction (or disjunction) of several literals, and **C** is the conclusion that A supports. Calvin uses its rules to form an argument (not a proof) for each element in **A**. From arguments in favor of **A**, Calvin creates an argument for **C**. The representation of the argument contains both the rule and the arguments for the antecedents. However, more convincing arguments against the conclusion may be found, and Calvin's belief in it overturned. This is the main distinction between an argumentation system and a classical first-order logic system.

Calvin's rules contain several additional elements: a quality rating, a guard, and a confidence template. The quality rating and confidence template are used to judge the relative and absolute strengths of arguments. Guards prevent the engine from building arguments using rules that are not applicable to the current case. For example, Calvin knows that snow cover is more likely if sample age is inversely correlated with elevation. This is based on the knowledge that snow cover blocks cosmic rays and more snow falls at higher elevations, but only makes sense for sample sets with large elevation ranges. Otherwise, random differences in the data might be interpreted as a meaningful correlation. The guard on this rule tells the engine to ignore the rule unless this precondition holds. Other argumentation systems typically do not require an explicit guard mechanism because they instead defeat rules explicitly (Farley 1997), (Morge and Mancarella 2007).

The antecedents in a rule describe templates for the evidence that will satisfy that rule argue for the rule's conclusion. These patterns define both what evidence is needed to satisfy the rule and where that evidence can be located: that is, whether to build an argument for a new conclusion or refer to the data input by the user.

The arguments for a conclusion **C** are a collection of trees constructed by Calvin's engine by unifying rules with evidence. Alternatively, each argument can be viewed as a tuple of the conclusion and support for the argument, as in the Logic of Argumentation of (Krause et al. 1995). The root of each tree in the collection is a rule whose conclusion is **C**, such as the rule **A**=>**C**. Each child of this root is one of the literals in **A** unified with evidence. This evidence may be either additional collections of argument trees or a reference to the input data. Calvin's backwards-chaining engine generally makes no distinction between negative and positive evidence. This is not a valid method in classical logic, where the knowledge that **A**=>**C** certainly does not imply that **~A**=>**~C**. However, Calvin's reasoning is intended to mimic that of experts, who are not necessarily logical. Experts not only apply rules in this negative fashion, they regard it as a sufficiently defensible practice that they discuss it in published reasoning. For example, (Jackson et al. 1997) includes the statement that, since there is no visual evidence of erosion, erosion is unlikely in the area under consideration.

Weighing Arguments

Some arguments carry greater weight than others, but precise comparisons between arguments are not always easy to perform. For example, some arguments for exhumation on a hypothetical moraine might be:

(1) This moraine has a flat crest, which is a visual sign of matrix erosion. Matrix erosion causes exhumation.

(2) This landform is a moraine, and moraines usually have a matrix, which is soft and erodes quickly. Matrix erosion causes exhumation.

(3) This landform has samples as old as 50ky, and various processes often disturb the surface and cause exhumation over such a long time period.

Clearly (1) and (2) are similar arguments, sharing the same root rule. Calvin would derive these arguments as a single tree with two branches. However (1) is a stronger argument for exhumation because it draws on empirical observations rather than general knowledge about moraines. This issue is often handled in argumentation systems by referring to the *specificity* of arguments, with more-specific arguments carrying more weight (Elvang-Gøransson et al. 1993). However, (3) seems to contradict this choice of weighting: although it refers to information that is specific to this landform, it seems weaker than (1). Furthermore, the relationship between (2) and (3) is surprisingly difficult to quantify. How, then, are we to judge the strengths of these three arguments in a way that preserves the intuitive relationships between them?

The central principal of Calvin's confidence system is that not only can specific *evidence* be trivial or critical, but the *knowledge* used to connect the evidence to the conclusion is also of variable quality. Defining confidence with two dimensions allows us to clarify why one argument is better than another: (1) uses high-quality evidence and high-quality knowledge. (2) uses high-quality knowledge but only moderate-quality evidence, and (3) uses high-quality evidence but low-quality knowledge. Separating the sources of confidence greatly enhances our understanding of the strengths of these three arguments. To instantiate this, Calvin represents confidence as a two-dimensional vector. One element of the vector is determined by which rules were used to form the argument, and the other is determined by how closely the observed situation matched those rules.

As part of our knowledge engineering process, we asked experts about the strength of their belief in their heuristics to determine the appropriate qualitative validity to assign to each rule. When Calvin unifies evidence with a rule, it creates a confidence vector for the rule's conclusion from the closeness of the current situation to the rule's threshold(s) (closer to thresholds gives less confidence) and the validity assigned to the rule. Calvin's engine uses this confidence vector to find an overall confidence in chains of arguments and in sets of argument trees.

Using Confidence To judge the strengths of the arguments it generates, Calvin manipulates confidence values in two distinct ways. The first operates along a single chain of reasoning: snow cover is more likely in cold areas; this area is cold because it is at high elevation. Intuitively, it makes sense to choose the validity of the least-valid rule for the overall conclusion: the chain is only as strong as its weakest link. Applicability is 'created' by the direct use of observed evidence. In this case, how high the sampled area is, compared to what elevation is usually cold, determines the applicability. A few rules lower or raise the applicability of knowledge passed through them when they are applied. This is to handle situations where an `observation` is not specific to the knowledge being applied, as in argument (2) at the beginning of this section.

The second and more-complicated use of confidence occurs when a number of different chains of reasoning are all applied to the same conclusion (because an argument is a *collection* of trees), e.g.: (a) erosion is more likely because the landform is old; (b) erosion is less likely because there is no visible sign of it. A chain of reasoning supporting the conclusion might have higher validity but lower applicability than a chain of reasoning refuting the conclusion. There are often several independent chains of reasoning both supporting and countering the conclusion, each with its own confidence level. Calvin, like many existing argumentation systems (Prakken 2005), assigns confidence in two stages, first locally up a single chain of reasoning and then globally across many chains of reasoning arguing about the same conclusion.

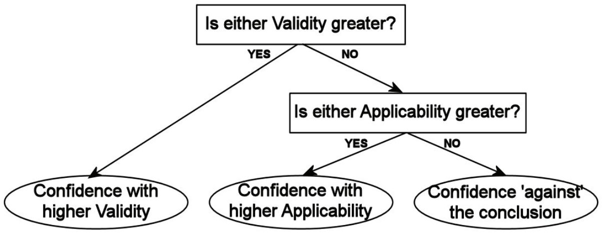

Figure 4: A decision tree for which confidence is greater in comparing opposing confidences.

To determine its overall confidence in a conclusion, Calvin first aggregates groups of lower-validity confidences in into higher-validity confidences. Then, if the highest-validity confidences for and against the conclusion are at least two levels apart, the highest-validity confidence is returned intact as the overall confidence: it is judged sufficiently strong to completely override the weaker rebutting evidence. A difference of two levels of validity implies a huge difference in overall confidence strength—it is the difference between a logical tautology and a statement such as 'frost heaving sometimes occurs in cold areas.' In contrast, a single level of difference in validity is less drastic, for example the difference between the preceding statement and a statement that 'snow cover is plausible in cold areas.' The resulting confidence in other situations is illustrated in Figure 4 and Table 1. Figure 4 indicates which confidence is considered greater and assigned to the overall conclusion. However, when the two confidences are close, Calvin reduces its overall confidence in the conclusion according to how close the two competing confidences are. Table 1 shows the possible ranks of confidence reduction and when they apply.

Table 1: Reduction Operations in Confidence Combination

Reduction Operation	Occurs When	
	Validity > AND	Validity = AND
Do Nothing	Applicability >>	
Applicability-	Applicability >=	Applicability >>
Applicability- -		'Against' Applicability >
Validity- -	Applicability <	
Validity-, Applicability-		'For' Applicability >
Validity-, Applicability-		Applicability =

Calvin, then, reproduces expert reasoning by considering a set list of hypotheses one at a time, creating arguments for and against each hypothesis. Evidence may take the form of a single comparison or a complete subargument. Calvin then weighs these arguments based on the quality of knowledge and certainty of evidence used to generate them. This weighting results in both absolute and relative judgments of argument strength, as well as indicating the strongest and weakest points of each argument.

Results

Experts publish some of their qualitative reasoning about a landform when they publish its age. While this presentation is usually incomplete, it typically includes information about both rejected and accepted conclusions. We used these to assess Calvin's ability to reproduce human expert reasoning. We compared Calvin's reasoning to the reasoning in eighteen randomly-selected papers discussing one or more isotope dating problems in detail. These publications provide a broad basis of comparison. To compare Calvin's output with this prose, we extracted every statement from these papers that made an assertion and distilled it to the conclusion being argued and the evidence presented for that conclusion. We then entered all the data given in the paper, ran Calvin, and compared its output to these argument summaries.

Calvin performed quite well at reproducing arguments published in isotope dating papers. For 62.7% of the arguments in published work, Calvin came to the same conclusion, supported by the same evidence, at about the same confidence level as the argument in the original paper (we found judging confidence levels from prose relatively difficult, and only divided these arguments into 'strong' and 'weak' categories). On a further 26.1% of these arguments, Calvin succeeded in two of these elements (recognizing the same evidence as important but to a different conclusion, coming to the same conclusion with different evidence or a vastly different confidence level). More detailed results are presented in (Rassbach 2009). In a few cases, Calvin produced arguments that did not appear at all in the original paper. In one such case, when examining (Ballantyne and Stone 1998), Calvin argued that the samples were exhumed. The main evidence for this is a disagreement with ages determined for this landform via other methods. To judge these results, we asked a domain expert to assess Calvin's new argument. He responded:

> I think I see both sides here. From the results, the fact that the ages are younger than the C14 data means that exhumation should be taken very seriously (...) there is not much in the way of material that could bury them. However the peaks themselves are eroding...

Clearly choosing not to explicitly address exhumation in (Ballantyne and Stone 1998) was a major oversight, given the amount of unclear and conflicting evidence that may or may not be indicative of it. Although Calvin does not give exactly the same argument, it has found a major gap in the reasoning published by these authors.

In some cases, Calvin produced arguments strikingly similar to the statements in the paper. These similarities were especially obvious when the authors of the paper expressed significant doubt about their conclusions. For example, consider this passage from (Briner et al. 2005):

> "The ca. 56 ka age on the Jago lateral moraine appears to be a clear outlier that we attribute to inheritance. The age of the Okpilak ridge is uncertain; correlation with the Jago ridge supports the suggestion that the two older boulders from the Okpilak ridge contain inherited isotopes. Alternatively, both ridges might be pre-late Wisconsin in age, and the young age cluster on the Jago ridge records accelerated moraine degradation and consequent boulder exhumation during the late Wisconsin. On the other hand, the stabilization age indicated by the (...) ca. 27 ka age is consistent with Hamilton's (1982) age constraints for deglaciation..."

Calvin finds it quite likely that the 56ka sample is an outlier and attributes the difference to inheritance. However it, too, grapples with explaining the age of the Okpalik ridge: inheritance is supported by correlation with the Jago moraine, the 25ka expected age, and the climate of the area. However, this implies that 2/3 of the samples from that ridge contain significant inheritance, leading to a conflicted overall argument for inheritance. Calvin also finds significant support for exhumation on both moraines, coming to the same uncertain conclusion as the authors.

Conclusion

Calvin is a fully implemented argumentation system in use by experts in cosmogenic isotope dating (it has been downloaded 178 times). Because of its nature as a concrete system, building Calvin required us to solve a complex problem: how to best describe and compare confidence. Our solution, a two-element vector to represent confidence, and the associated system for weighing rebutting arguments appears to be novel. This system, while complex in implementation, elegantly captures the argument comparisons we observed experts making.

Calvin is an argumentation system because our goal was to reproduce the structure of expert reasoning. Although isotope dating experts may speak in terms of probabilities and chains of reasoning, they, like most scientists, do not reason in a probabilistically or logically correct manner. Thus, an inflexible system of probabilities or logic would find it difficult to accurately reproduce the reasoning of experts in this field. Expert argument comparisons more closely resemble possibilistic logic (Dubois and Fagier 2005), (Farreny and Prade 1996) and our own confidence system than either a Bayesian or pure logic system.

While Calvin's initial results are extremely promising, we are in the process of planning a more rigorous study (with automatic annotation (White 2009)) to more completely test its success at solving this problem. Furthermore, we believe that Calvin's confidence system will translate well to other problems where weighing competing arguments is difficult—both in other scientific fields such as forensic linguistics and problems in other domains, such as the game of bridge. We hope to identify if there is a cognitive mechanism that weighs rebutting arguments in a consistent way across domains and, if so, to elucidate that mechanism.

References

Anderson, K., Bradley, L., Zreda, M., Rassbach, L., Zweck, C., and Sheehan, E. 2007. ACE: Age Calculation Engine—A Design Environment for Cosmogenic Dating Techniques. In Proceedings of the First International Conference on Advanced Engineering Computing and Applications in Sciences. Tahiti.

Ballantyne, C.K., Stone, J.O., and Fifield, L.K. 1998. Cosmogenic Cl-36 dating of postglacial landsliding at The Storr, Isle of Skye, Scotland. The Holocene 8(3): 347-351.

Briner, J.P., Kaufman, D.S., Manley, W.F., Finkel, R.C., and Caffee, M.W. 2005a. Cosmogenic exposure dating of late Pleistocene moraine stabilization in Alaska. GSA Bulletin 117(7/8): 1108-1120.

Chamberlain, T. C. 1965. The method of multiple working hypotheses. (Reprint of 1890 Science article.) Science 148: 754-759.

Dubois, D. and Fargier, H. 2005. On the Qualitative Comparison of Sets of Positive and Negative Affects. Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Lecture Notes in Computer Science 3571: 305-316.

Elvang-Gøransson, M., Krause, P., and Fox, J. 1993. Acceptability of Arguments as "Logical Uncertainty." In Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty. Berlin, Germany.

Farley, A.M. 1997. Qualitative Argumentation. In Proceedings of the Eleventh International Workshop on Qualitative Reasoning. Cortona, Italy.

Farreny, H. and Prade, H. 1986. Default and inexact reasoning with possibility degrees. IEEE Transmissions on Systems, Man and Cybernetics 16(2): 270-276.

Jackson Jr., L.E., Phillips, F.M., Shimamura, K., and Little, E.C. 1997. Cosmogenic 36Cl dating of the Foothills erratics train, Alberta, Canada. Geology 25(3): 195-198.

Krause, P., Ambler, S., Elvang-Gøransson, M., and Fox, J. 1995. A Logic of Argumentation for Reasoning Under Uncertainty. Computational Intelligence 11: 113-131.

Morge, M. and Mancarella, P. 2007. The Hedgehog and the Fox: An Argumentation-Based Decision Support System. In Proceedings of the 4th International Workshop on Argumentation in Multi-Agent Systems.

Prakken, H. 2005. A study of accrual of arguments, with applications to evidential reasoning. In Proceedings of the Tenth International Conference on Artificial Intelligence and Law, 85-94. New York.

Putkonen, J. and Swanson, T. 2003. Accuracy of cosmogenic ages for moraines. Quaternary Research 59: 255-261.

Rassbach, L. 2009. Calvin: Producing Expert Arguments about Geological History. Ph.D. diss., Dept. of Computer Science, Univ. of Colorado, Boulder, Colo.

White, E. 2009. Pattern-Based Recovery of Argumentation from Scientific Text. Ph.D. diss., Dept. of Computer Science, Univ. of Colorado, Boulder, Colo.