

Argumentation schemes: From genetics to international relations to environmental science policy to AI ethics

Nancy L. Green

University of North Carolina Greensboro, Greensboro, NC 27402, USA

E-mail: nlgreen@uncg.edu

Abstract. Argumentation schemes have played a key role in our research projects on computational models of natural argument over the last decade. The catalogue of schemes in Walton, Reed and Macagno’s 2008 book, *Argumentation Schemes*, served as our starting point for analysis of the naturally occurring arguments in written text, i.e., text in different genres having different types of author, audience, and subject domain (genetics, international relations, environmental science policy, AI ethics), for different argument goals, and for different possible future applications. We would often first attempt to analyze the arguments in our corpora in terms of those schemes, then adapt schemes as needed for the goals of the project, and in some cases implement them for use in computational models. Among computational researchers, the main interest in argumentation schemes has been for use in argument mining by applying machine learning methods to existing argument corpora. In contrast, a primary goal of our research has been to learn more about written arguments themselves in various contemporary fields. Our approach has been to manually analyze semantics, discourse structure, argumentation, and rhetoric in texts. Another goal has been to create sharable digital corpora containing the results of our studies. Our approach has been to define argument schemes for use by human corpus annotators or for use in logic programs for argument mining. The third goal is to design useful computer applications based upon our studies, such as argument diagramming systems that provide argument schemes as building blocks. This paper describes each of the various projects: the methods, the argument schemes that were identified, and how they were used. Then a synthesis of the results is given with a discussion of open issues.

Keywords: Argumentation schemes, computational models of natural argument, argument corpus, argument generation, argument mining, argument modeling systems

1. Introduction

Argumentation (or argument) schemes [60] have played a key role in our research projects on computational models of natural argument over the last decade. Argumentation schemes have been described as patterns of acceptable, presumptive arguments in law, science and everyday conversation. The Walton et al. catalogue of schemes served as our starting point for analysis of the naturally occurring¹ arguments in written text, i.e., text in different genres having different types of author, audience, and subject domain (genetics, international relations, environmental science policy, AI ethics), for different argument goals, and for different possible future applications (Table 1). We would often first attempt to analyze the arguments in our corpora in terms of those schemes, then adapt schemes as needed for the goals of the project, and in some cases implement them for use in computational models.

¹“Naturally occurring”, as opposed to arguments created by research participants or by students in response to a school assignment.

Table 1
Research projects involving argument schemes

Genre	Author	Audience	Argument goal	Domain	Possible application	Project references
Genetic counseling patient letter	Genetic counselor	Client, Healthcare providers	Support clinic's conclusions; Emotional support to client	Clinical genetics	Letter generation (GenIE Assistant)	[14–17,32,33]
Genetic counseling case study	Genetic counselor	Student	Support clinic's conclusions	Clinical genetics	Argument modeling system (GAIL)	[22,34]
Genetic testing promotional brochure	Genetic testing company	Patient	Persuade patient to seek genetic testing	Clinical genetics	Healthcare consumer information system	[18]
Genetics research journal article	Geneticist	Others in same field	Support new knowledge claims	Genetics	Argument mining	[19–21,23,24,27]
International relations journal article	International relations analyst	Analyst or student	Support analysis and policy recommendations	International Relations	Argument modeling system (AVIZE)	[25,29]
Science policy journal article	Scientist or science writer	Science-literate reader	Support policy recommendations	Environmental science	Rhetoric and argument modeling system	[26,28]
AI ethics case studies	Ethics expert	Student	Support ethical acceptability of agent's action	AI Ethics in military and healthcare applications	Argument modeling system (AIED)	[31]

Among computational researchers, the main interest in argumentation schemes has been for use in argument mining by machine learning (ML) methods [41,56]. Feng and Hirst [13] proposed that after an argument's premises and conclusion had been identified automatically, recognition of its argumentation scheme could be used to infer implicit elements of the arguments (enthymemes). They created ML classifiers to recognize several argument schemes of the Walton et al. catalogue in the Araucaria corpus, which contains annotated premises and conclusions of arguments from newspaper articles and court cases [52]. Lawrence and Reed [40] experimented with a corpus of arguments extracted from a 19th century philosophy text, in which proposition types had been identified by ML. Then groups of propositions that could belong to the same argumentation scheme (several schemes from the Walton et al. catalogue) were recognized as arguments, where missing elements of a scheme were assumed to indicate enthymemes.

In contrast, a primary goal of our research has been to learn more about written arguments themselves in various contemporary fields. Our approach has been to manually analyze semantics, discourse structure, argumentation, and rhetoric² in texts. Another goal has been to create sharable digital corpora containing the results of our studies. For one, an argument-annotated corpus from the natural sciences literature would be a valuable resource since the lack of such corpora is a major obstacle to research.

²However, we did not analyze rhetoric systematically until we read the 2017 8(3) special issue of *Argumentation and Computation on rhetoric*.

Our approach has been to define argument schemes for use by human corpus annotators or for use in logic programs for argument mining. The third goal is to design useful computer applications based upon our studies. Currently the field of genetics is highly active with a large and rapidly growing body of research that could benefit from patient-centered applications, such as natural language generation of genetic counseling materials, as well as intelligent search and summarization tools for genetics researchers. Also, we have been interested in designing argument diagramming applications that provide argument schemes as building blocks, e.g., applications to help students and the public to become wiser consumers of arguments in the field of environmental science policy, or to help computer science students to be conscious of issues in AI ethics.

This paper is organized as follows. First, the various projects are described, more or less chronologically, in sufficient detail to describe the methods, the argument schemes that were identified, and how they were used. Then a synthesis of the results is given with a discussion of open issues.

2. Projects

2.1. GenIE (*Genetics Information Expression*) assistant

The GenIE Assistant [32] was a prototype system for generation of the first draft of genetic counseling letters. It was implemented as a testbed for research on natural language generation of transparent biomedical argumentation. The system architecture included (1) a knowledge base (KB), a causal network representation of domain knowledge about genetic disorders and of information about a patient's case; (2) a discourse grammar with genre-specific rules for creating an abstract representation (DPlan) of a letter to a client; (3) an argument generator which generated arguments in propositional form for claims in the DPlan; (4) an argument presenter which made changes in the DPlan for the sake of text coherence and transparency; and finally, (5) a linguistic realization component to render the DPlan as English text. The DPlan contained propositions about the patient's case from the KB and was structured by text coherence relations of Rhetorical Structure Theory (RST) [43].

GenIE's argument generator created arguments for claims in the DPlan using its argument schemes and information in the KB. The argument schemes and the format of the KB were informed by a corpus study of genetic counseling letters.³ The corpus study [14] revealed a conceptual model of genetics used in communication with clients. The conceptual model consisted of general concepts, such as *history*, *symptom*, *genotype*, and *test result*, and causal relations between concepts, which we described in terms of qualitative constraints of qualitative probabilistic⁴ networks [10]. The KB could be instantiated for different genetic disorders and for different patients' cases.

Rather than refer to domain content, the argument schemes were formulated at a higher level of abstraction in terms of qualitative network node variables and relations. For example, to paraphrase one of GenIE's Effect to Cause schemes shown in Fig. 1, the claim that a node A in the KB was at or above a certain threshold is supported by the Data (premise) that a node B in the KB is at or above a certain threshold, and by the Warrant (premise) that there is a positive influence relation from A to

³Nine letters ranging in length from 446 to 1537 words on seven genetic conditions, representing the three main types of single-gene inheritance (recessive, dominant, new mutation), written by four genetics counselors from different institutions. Although small in comparison to corpora used for machine learning, the corpus was representative of the genre and contained a variety of argument patterns.

⁴Although not incorporated into the final GenIE system, we also analyzed the different uses of probability statements in arguments in the corpus [15,16].

Claim: $A \geq a$	<i>P has 2 mutated GJB2 alleles</i>
Data: $B \geq b$	<i>P has hearing loss</i>
Warrant: $S'(\langle A, a \rangle, \langle B, b \rangle)$	<i>Having 2 mutated GJB2 alleles can result in ... hearing loss</i>
Applicability Constraint: $\neg \exists C X ((C, A) \langle B, b \rangle): C \geq c$	<i>Unless something else is thought to be the cause</i>

Fig. 1. Simple effect to cause scheme in GenIE assistant (example in italics).

Universal condition:

- Data: Patient inherited genetic condition C from both parents.
- Warrant: "It is estimated that everyone has a number of changes that can cause problems in our children if our partner is also a carrier."
- Conclusion (implicit): You are not to blame for patient's condition C.

Fig. 2. Universal condition affective argument scheme.

B. Such a scheme could be used to generate an argument that a certain patient has a genotype of two mutated alleles of a GJB2 gene based on the data that he has hearing loss, and the warrant that this genotype can result in hearing loss. Note that, unlike the schemes in the Walton et al. catalogue [60], the GenIE schemes distinguished premises as Data or Warrant [57]. The Data premise represented patient-case-specific information (stored in nodes of the KB), while the Warrant represented general biomedical principles (represented as qualitative constraints in the KB). The Data/Warrant distinction played a role in the organization of the text (DPlan) and in decisions of the argument presenter. In the DPlan, the Warrant was represented as the satellite of an RST Background relation, whose nucleus was an RST Evidence relation; the satellite of the Evidence relation was the Data, and the nucleus was the Claim [17]. The argument presenter applied heuristics to the DPlan to make a letter more concise. One heuristic was if two adjacent arguments in the DPlan included the same warrant, the warrant in the second argument was replaced by an adverb such as 'similarly'.

GenIE's schemes also included "applicability constraints", which were used to determine the applicability of a scheme during the argument generation process. In Fig. 1, the Effect to Cause scheme included an applicability constraint that can be paraphrased as "there is no other node in the KB which has positive influence on B." The applicability constraints were used as critical questions in an interactive version of GenIE.

In subsequent work we identified several affective argument schemes in the corpus for mitigating the client's possible negative reaction to information presented in the letters [33]. For example, see Fig. 2. If the partially generated letter contained the information in Data, then the Warrant would be inserted into the letter. The phrasing of the sentence expressing the Warrant was derived from the corpus. The Conclusion of the scheme was not explicitly stated in the generated letter, consistent with use of this scheme in the corpus.

2.2. GAIL (Genetics Argument Inquiry Learning) system

The argument generator and schemes developed for GenIE were repurposed as the core of an educational argument modeling system for biology students, GAIL [22,34]. GAIL's user interface presented the student with information to use in constructing graphical arguments: the Problem (to give an argument for a certain claim), the Data (information from a fictitious patient's medical record), several possible Hypotheses, and Connections (possibly relevant principles of genetics, i.e. warrants). See a screenshot in Fig. 3. For example, a Problem was "Give two arguments for the hypothesis that the patient, J. B., has cystic fibrosis (has two mutated copies of the CFTR gene)." The Data about the patient included items such as "History of respiratory problems. During her second year, J.B. developed a

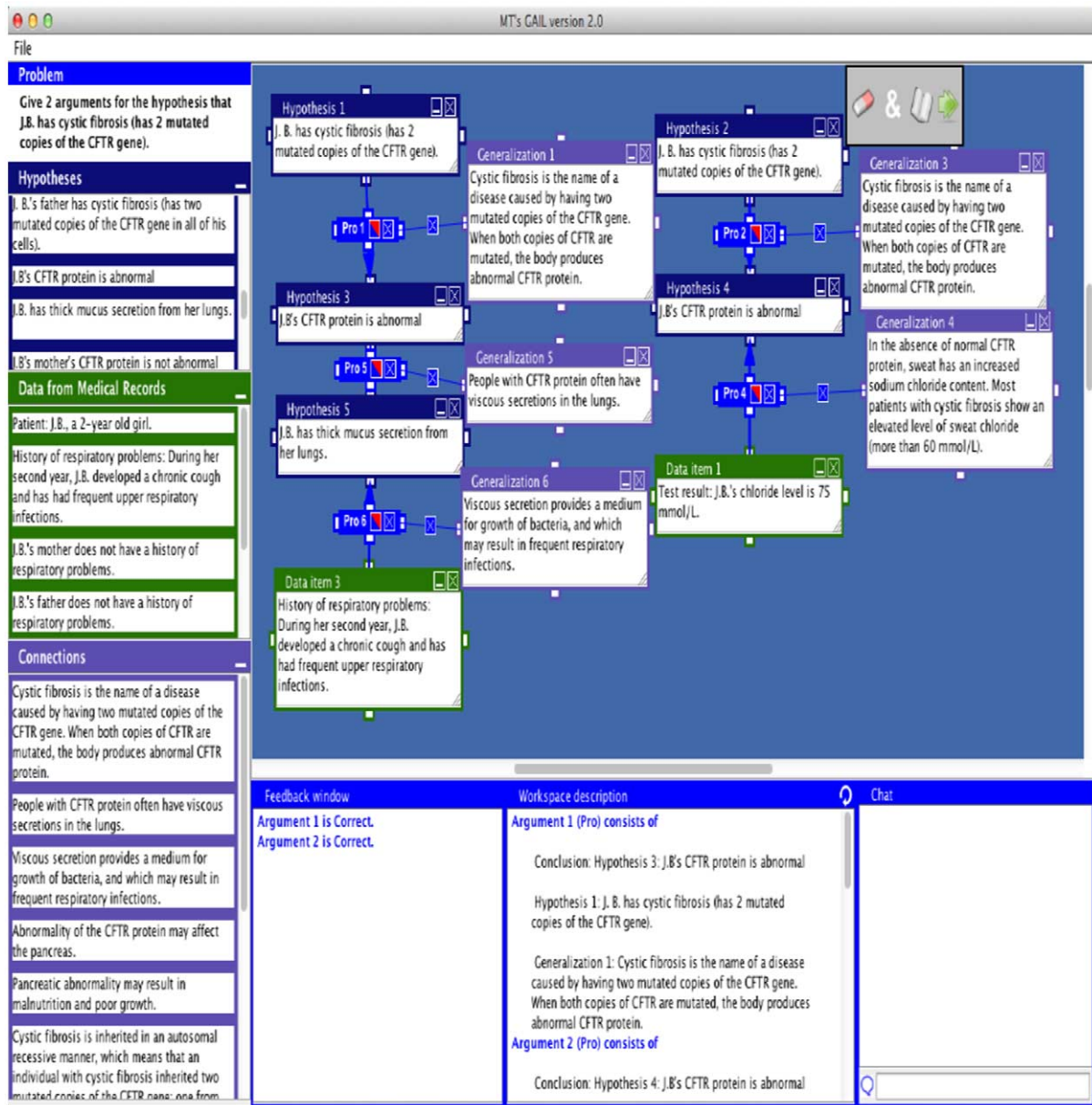


Fig. 3. GAIL screenshot showing two arguments. (Main claim at top of each chain; warrants at right-angles to chain.)

chronic cough and has frequent upper respiratory infections”, and “J.B.’s mother does not have a history of respiratory infections”. The Connections included items such as “When both copies of CFTR are mutated, the body produces abnormal CFTR protein”, “Abnormality of the CFTR protein may affect the pancreas”, “People with CFTR protein often have viscous secretions in the lungs”, etc. To construct an argument, the learner could select appropriate items from the Data and Connections lists, drag them into the argument diagramming area, and connect them into a Toulmin-style (Claim-Data-Warrant) diagram of an argument (or chain of arguments as shown in Fig. 3).

An authoring tool enabled instructors to create a knowledge base (KB) in the same format as used in GenIE to represent domain knowledge and information about a patient's case. The authoring tool also created a mapping from the English text to be presented to the student on the screen to the underlying KB. After a student constructed graphical arguments for a claim, GAIL mapped the student's arguments to their internal representations. Then GAIL used GenIE's argument generator to create expert arguments for the problem. By comparing the student's arguments to GenIE's arguments, GAIL was able to generate feedback on the structure and content of the student's arguments. Previous educational argument modeling systems required manual authoring of expert arguments in order to enable generation of feedback on content. Also, although not implemented in GAIL, it was noted that use of the critical questions of the argument schemes could support generation of additional feedback, such as "Can you make an argument for a diagnosis other than cystic fibrosis that explains the patient's malnutrition as well as his frequent respiratory infections?".

2.3. Analysis of genetics testing company promotional brochure

In this project we analyzed written arguments and persuasive visual devices in a five-page patient brochure published by a genetic testing company [18]. The goal of the brochure was to motivate patients to seek genetic testing from that company. We identified variants of causal and practical reasoning arguments and a specialization of fear appeal arguments based upon Protection Motivation Theory [49].⁵ We noted that a patient might lack quantitative skills necessary for understanding probability statements used in arguments. Furthermore, we showed that while the brochure addressed critical questions that would support its arguments, analysis of two other sources revealed answers to critical questions that challenged the brochure's arguments. In addition, the brochure's arguments could be challenged by re-framing, qualifying, or disputing elements of its arguments with information from the other sources.

2.4. Towards argument mining genetics research articles

After considering arguments addressed to the lay reader in the GenIE project, we investigated how argumentation is used in genetics research articles. Such articles are written by and for other scientists. Whereas the genetic counseling letters conveyed a simplified, accepted model of genetics to warrant the claims of the letters, the goal of scientific research is to discover new knowledge by rejecting or refining current models or proposing a new model. Thus, the argument schemes used in GenIE were not sufficient to model scientific research.

As a step towards creating an argument-annotated corpus of genetics research articles [19], we analyzed some of the arguments in a representative article [55]. The analysis distinguished premises as Data and Warrant. It found that some missing (implicit) Warrants were necessary for argument acceptability.⁶ Also, conclusions of arguments were sometimes implicit; and some conclusions (explicit or implicit) functioned as implicit premises of subsequent arguments. In follow-up work [21], we defined ten causal argument schemes based on analysis of that article and three other genetics articles [6,9,44], e.g., Effect to Cause, Failed to Observe Effect of Hypothesized Cause, Consistent with Predicted Effect, Hypothesize Candidates, Eliminate Candidates, and Joint Method of Agreement and Difference. The argument scheme definitions were stated in general terms rather than in terms of genetics concepts.⁷ Note that

⁵When not otherwise noted, in this paper lower-case initial schemes refer to schemes in the Walton et al. catalogue [60].

⁶Our analysis of the arguments was later confirmed by a domain expert.

⁷Also, in these definitions the premises did not distinguish Data and Warrant.

Excerpt from [58] with annotated spans:

Given our interest in human neurological disease we sought to identify
 any cognate human disorders where linkage had been established to the syntenic region
 of the human genome,
 but where no causal mutation had been identified.
 SCA15, an adult-onset autosomal dominant progressive ataxia is linked to this locus [5].

 Although missense mutation of ITPR1 had previously been ruled out [2]
 and the mode of inheritance was inconsistent with that seen in the Itpr1Δ18 and Itpr1opt
 mice,
 the phenotypic presence of ataxia in the mice
 led us to reexamine this candidate gene as a possible cause of SCA15.

Annotated argument:

```
<argument ID="argument2" scheme="Analogy">
<premise ID="premise1" span="(1,2,5)" paraphrase="The movement disorder in the mice is similar to
ataxia in humans">
<premise ID="premise2" implicit="yes" conclusion_of="argument1" paraphrase="A mutation of Itpr1
may be the cause of the ataxia-like phenotype of the mice">
<premise ID="premise3" span="1" paraphrase="mouse Itpr1 is syntenically related to human ITPR1">
<conclusion span="6" paraphrase="a mutation in ITPR1 may be a cause of ataxia in humans">
<critical_question ID="cq1" span="3" paraphrase="Is the difference between missense and deletion
mutation of Itpr1/ITPR1 significant?">
<critical_question ID="cq2" span="4" paraphrase="Is the difference between mode of inheritance of
mouse and human ataxia significant?">
</argument>
```

Fig. 4. Example of proposed annotation of article in [20].

Effect to Cause, Failed to Observe Effect of Hypothesized Cause, and Eliminate Candidates are similar to the abductive argumentation scheme, argument from falsification, and argument from alternatives, respectively, in the Walton et al. catalogue [60]. Joint Method of Agreement and Difference was based on Mill's Method of Agreement and Difference [37].

A pilot study [21] on using our proposed scheme definitions to analyze arguments in genetics research articles found that undergraduate biology students had considerable difficulty applying the schemes correctly.⁸ However, a smaller follow-up study with faculty as annotators found that they could apply the schemes successfully. Later, other researchers [47] independently found the schemes applicable to the analysis of five biochemistry research articles and speculated that the schemes may be applicable to the experimental biomedical literature in general.

Next, in order to create a freely available argument-annotated corpus, we selected a genetics research article [58] from the CRAFT corpus (CRAFT 17590087) to be annotated.⁹ In [20] we proposed an annotation system that would specify where in the text an argument's premises, conclusion, and, possibly, answers to critical questions are expressed. The top part of Fig. 4 shows the proposed annotation of spans of text in an excerpt from the article. The bottom part shows how that excerpt was analyzed as a kind of causal argument from analogy. The first premise was shown as coming from spans 1, 2, and 5 (as indicated at the top of the figure), and the second premise was analyzed as implicit. Spans 3 and 4 were analyzed as responses to two critical questions of the scheme.

Note that the definitions of argument schemes in [21] did not refer to genetics concepts. However, given the results of the pilot study with undergraduates, we wanted to make the task easier for future annotators. Thus we redefined the argument schemes in terms of a small set of domain concepts such as genotype and phenotype, e.g. as shown in Fig. 5. The revised annotation guidelines¹⁰ include a tax-

⁸Factors that may have contributed to this difficulty include lack of motivation (the students did not volunteer for the task, which was given to them during their regular laboratory meeting) and lack of experience reading research articles.

⁹Articles in the CRAFT corpus [5,59] may be redistributed. The corpus includes annotations of concepts and linguistic structure.

¹⁰Available at <https://github.com/greennl/BIO-Arg>.

Scheme: Method of Agreement
Premise: A group of individuals I have atypical phenotype P
Premise: All of the individuals in I have atypical genotype M.
Conclusion: M may be the cause of P (in I).

Fig. 5. Genetics-specific method of agreement argument scheme.

- 1 Causation
 - 1.1 One Group
 - 1.1.1 Agreement Arguments
 - 1.1.1.1 Agreement
 - 1.1.1.2 Failed Method of Agreement (effect)
 - 1.1.1.3 Failed Method of Agreement (cause)
 - 1.1.2 Eliminate Candidates
 - 1.1.3 Explanation-Based
 - 1.1.3.1 Effect to Cause
 - 1.1.3.2 No Effect to No Cause
 - 1.1.3.3 Consistent with Predicted Effect
 - 1.2 Two Group
 - 1.2.1 Difference
 - 1.2.2 Analogy (Causal)
 - 1.2.3 Explanation-Based
 - 1.2.3.1 Consistent Explanation
 - 1.2.3.2 Difference Consistent Explanation
- 2 Other
 - 2.1 Classification
 - 2.2 Confirmation

Fig. 6. Taxonomy of argument schemes in genetics research articles from annotation guide.

```
arg(
  scheme('Agreement'),
  premise(have_phenotype(G, P)),
  premise(have_genotype(G, M)),
  conclusion(cause(M, P)))
:-
  group(G),
  have_phenotype(G, P),
  have_genotype(G, M).
```

Fig. 7. Prolog rule for extracting argument based on method of agreement.

onomy of the argument schemes as an aid to annotators (Fig. 6). The causal schemes are divided based on whether they involve one or two groups of individuals. The One Group branch includes argument schemes related to Effect to Cause and Mill's Method of Agreement. The Two Group branch includes schemes related to Mill's Method of Difference and a variant of argument from analogy with a causal conclusion. Both branches of the taxonomy include arguments based on consistency.

After revising the scheme definitions in terms of genetics concepts and relations,¹¹ we envisioned how the argument schemes could be used for argument mining if implemented as logic programming rules in Prolog [23]. We defined seven argument schemes as Prolog rules based on analysis of the first eight paragraphs of the ten paragraph Results/Discussion section of the article selected for annotation [58]. For example, an argument scheme based on Method of Agreement could be implemented for this domain as shown in Fig. 7. The rule says that an argument (whose scheme name is 'Agreement') may consist of the premise that a group of individuals G has abnormal phenotype P, the premise that G has abnormal genotype M, and the conclusion that M may be the cause of P, provided that G is a group such that G's phenotype is P and G's genotype is M. Note that by implementing the scheme in this way, it can be used

¹¹Unfortunately we never received extramural funding to support creation of an argument-annotated corpus of genetics research articles. Thus, we refocused on the argument mining research described in the rest of this section.


```

<entities-props>
<entity ID="group1" paraphrase="the affected mice" />
<entity ID="pheno1" paraphrase="ataxia-like movement disorder" />
<prop>have_pheno(group1, pheno1)</prop>
<entity ID="geno1" paraphrase="homozygous mutation on chromosome 6qE1" />
<prop>have_genotype(group1,geno1)</prop>
</entities-props>

<argument scheme="Agreement">
<premise-list>
<premise prop="have_pheno(group1, pheno1)" />
<premise prop="have_genotype(group1,geno1)" />
</premise-list>
<conclusion inferred-prop="cause(geno1,pheno1,group1)"
paraphrase="A homozygous mutation on chromosome 6qE1
may be the cause of the ataxia-like disorder in the affected mice" />
</argument>

```

Fig. 8. Annotation of some entities and relations, and an argument.

to extract¹² an argument whether the conclusion is explicit or implicit in the text.¹³ Also, by providing the name of the argument scheme, it is possible to associate critical questions with an argument.

As described in [23], to make use of the rules to extract arguments, first, BioNLP tools [51] would be used to extract relations from a text for the predicates in the argument schemes, such as *have_phenotype*. After the relations had been extracted, the argument schemes could be applied to the extracted relations to recognize the arguments in the text. To demonstrate this approach, we manually annotated the entities, relations and arguments in the article as described in [24].¹⁴ An excerpt of the annotated article is shown in Fig. 8. Entities and propositions are listed immediately following the discourse segment in which they are expressed. The arguments are listed immediately following the entities and propositions that are used in the arguments. (Unlike the proposed annotation system in [20], we did not annotate which span(s) expressed a premise or conclusion.) In Fig. 8, three entities have been identified in the preceding discourse segment (not shown in the figure): group1, pheno1, and geno1. Identifiers were assigned by the annotator; paraphrases are optional documentation provided by the annotator. As shown at the top of the figure, two propositions have also been identified in the preceding segment. Following that, the annotation of an argument provides the premises and conclusion of an argument that has been identified in the segment. The argument schemes implemented as rules were tested to confirm that they could extract the annotated arguments.

In [23] we suggested that in the future argument schemes might be acquired by inductive logic programming [48] from articles whose entities, relations, and arguments had been similarly annotated. We are currently experimenting with use of inductive logic programming to derive the argument schemes from the annotated article [27].

2.5. Argument Visualization and Evaluation (AVIZE) in international relations

In our next project, we approached a very different domain from the biology-related domains of our previous work. We found that international relations arguments involve the beliefs, goals, appraisals, actions and plans of social actors such as countries, governments, and politicians. The goal of this project

¹²Although the rule can be said to extract the arguments in the sense of finding arguments in a text, in another sense, it is generating arguments like the argument generator did in GenIE. This is different from the sense in which arguments are extracted by current machine learning approaches to argument mining.

¹³In fact, due to the flexibility of Prolog, the rules could be used to restrict the arguments that are returned by partially instantiating the rule, e.g., to find all arguments such that some genotype causes a certain phenotype.

¹⁴The annotated article is available at <https://github.com/greennl/BIO-Arg>.

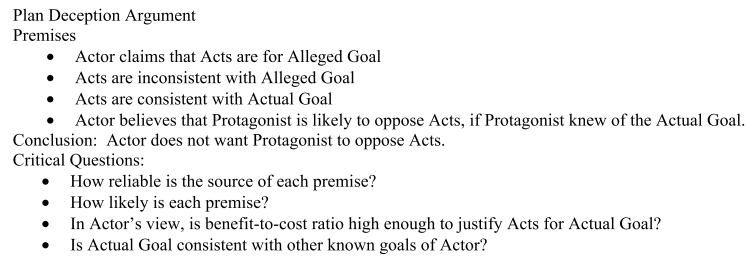


Fig. 9. Plan deception scheme in AVIZE.

was to develop a prototype argument modeling tool, AVIZE (Argument Visualization and Evaluation), to assist in the construction and (human) evaluation of arguments in this domain, based upon evidence of varying plausibility collected from sources of varying reliability. In addition to diagramming tools, AVIZE came with a set of argument schemes that we identified by analysis of several articles written by international relations experts.¹⁵ For example, the Plan Deception scheme is shown in Fig. 9.

We identified the following other schemes used to argue about an Actor's intentions: Plan Distraction, Inferred Plan, Coercion, Increasing Boldness, Behavior Pattern, Inferred Positive Appraisal, Inferred Negative Appraisal. Like the Plan Deception scheme, these schemes are domain-specific but are related to AI plan recognition heuristics. In addition, we identified the following schemes for arguing in favor of a planned action of the protagonist: Practical Reasoning, Resist Coercion, and Avoid Negative Consequences.¹⁶ These latter schemes are related to more abstract schemes such as practical reasoning, argument from threat, and argument from negative consequences, respectively. (In subsequent work [25] in this domain, we proposed an Intentional Cause to Effect argument scheme and its critical questions for reasoning about future events resulting from an agent's intentional acts.)

Challenges to arguments were represented in several ways in AVIZE. First, the user could attach both supporting and opposing pieces of data to the premise of an argument and decide which to credit. In this application, it was possible for sources to conflict and to have varying credibility. As shown in Fig. 10, for example, the premise that Martians have landed on Earth was supported by a news item that hikers had found a Martian space ship in the Nevada desert but challenged by two other items attached to that premise. Second, AVIZE enabled the user to attach counterarguments and provide answers to critical questions challenging a claim (Fig. 11). Also, Fig. 11 shows that users could color-code claims and premises to show their confidence in a premise or conclusion.

2.6. *Analysis of argument and rhetoric in environmental science policy arguments*

In our next project, we analyzed uses of rhetoric in a genre combining scientific and policy arguments in two journal articles in the domain of environmental science [38,39]. The long-term goal was to build a rhetorically-annotated digital corpus for research on rhetorical persuasion in argumentation. The short-term goal was to inform the design of tools to help students analyze use of rhetoric in science policy arguments. The overarching arguments in the articles were instances of value-based practical reason-

¹⁵The schemes were based on an in-depth analysis of a 33-paragraph article [62] and a reading of several others.

¹⁶A kind of summary can be provided by listing the names of the schemes that were used from beginning to end of the Weinberger [62] article: Plan Distraction, Coercion, critical question of Coercion, Resist Coercion, Plan Deception, Inferred Plan, Coercion, Increasing Boldness, Coercion, Inferred Plan, Inferred Plan, Resist Coercion, Practical Reasoning, Avoid Negative Consequences, Avoid Negative Consequences, Practical Reasoning.

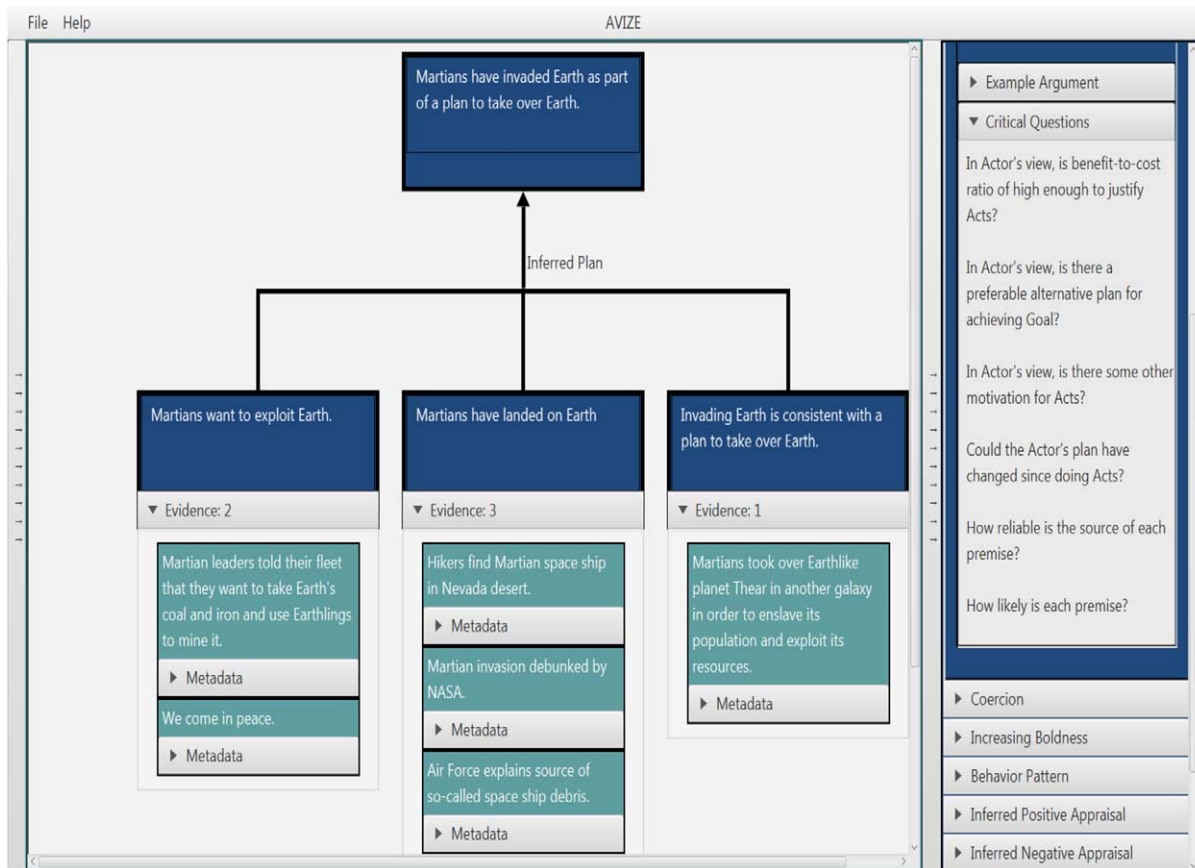


Fig. 10. AVIZE screenshot of inferred plan argument with evidence for each premise.

ing (VBPR), including a variant (VBPR-Knowledge Precondition) which argues for the need to acquire information as a precondition for action. In addition to the main VBPR arguments, in one article we noticed use of enthymematic VBPR arguments in the introduction to signal the authors' values and position on climate change. The articles also used two "rhetorical" argument strategies [45]: protocallepsis ("the rebuttal or refutation of anticipated arguments" (p. 241) and prolepsis as presage (warning of environmental catastrophe as motivation for action).

The analysis of rhetorical figures in the two articles was based on descriptions of figures in [11,35,36]. We found instances of a variety of types in each class of rhetorical figure as defined in [35]: schemes¹⁷ (phonetic, lexical or syntactic pattern), tropes (figures involving semantics such as metaphor), chroma (figures involving pragmatics such as rhetorical questions), and moves (discourse patterns such as protocallepsis). While our study did not reveal an association between rhetorical devices and argumentation schemes, it suggested that certain devices might play a role in automatic detection of rebuttals, e.g., rhetorical questions, sarcasm, satire, attacking the holder of an opposing view, concession, and conciliatio (identifying with certain concerns). Currently, we are studying the argumentative role of the rhetorical figure of antithesis in five environmental science articles [28].

¹⁷Not to be confused with argumentation schemes.

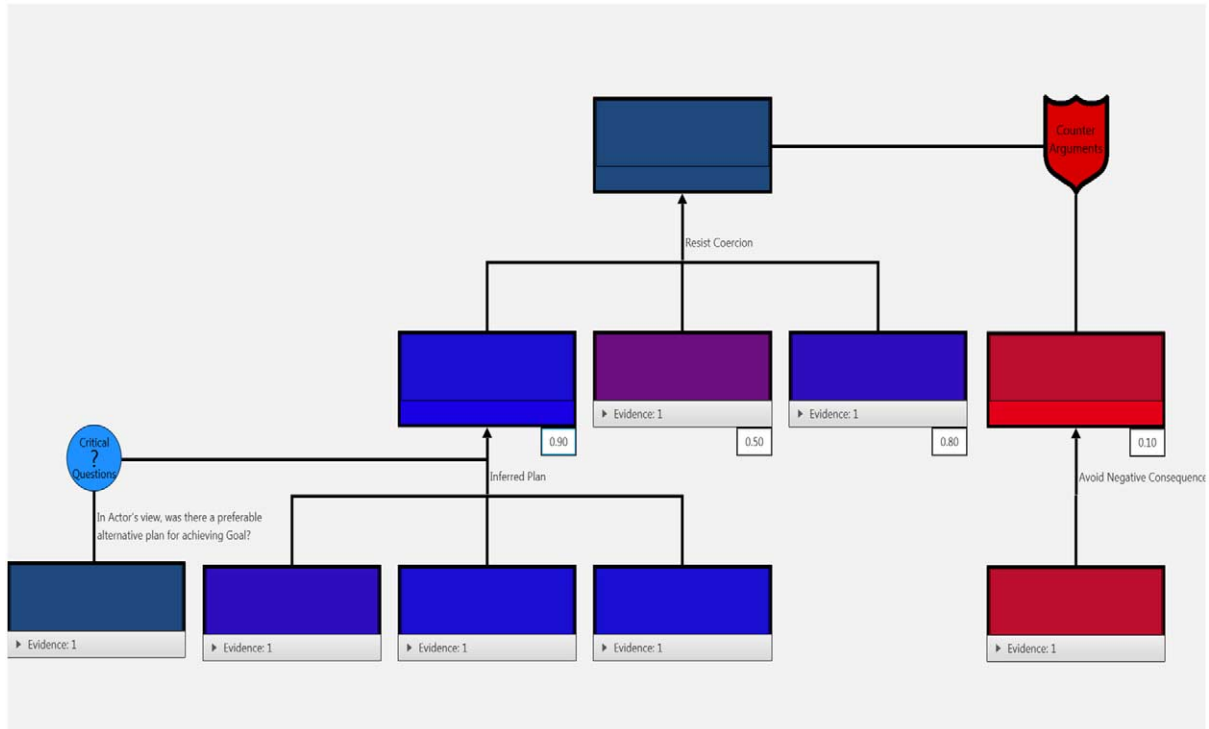


Fig. 11. AVIZE screenshot of arguments supporting and attacking a claim.

2.7. AI ethics debate (AIED)

In an AI Ethics¹⁸ course for computer science students, we focused on the design of explicit ethical agents [46]. An explicit ethical agent is an artificial agent that reasons about the ethical acceptability of its action using an explicit representation of ethical principles. In contrast, an implicit ethical agent is programmed so that its actions are consistent with human ethical judgments but it has no explicit representation of ethics. According to some AI researchers [1,54], it is preferable to develop explicit ethical agents since an autonomous agent may encounter situations requiring ethical decision making that were not anticipated by the agent's creators. Moreover, it is possible to examine the ethical justification for an explicit ethical agent's actions. (It is assumed that moral responsibility lies with the humans involved in the agent's creation or use: programmer, designer, purchaser, user, etc.). As a pedagogical tool, we developed several argument schemes for modeling the ethical acceptability of an explicit ethical agent's action in military and healthcare applications, incorporating issues that ethicists have raised for these domains.

Arkin [4] has done extensive research on design and implementation of explicit ethical agents for autonomous machines capable of lethal force, e.g., autonomous tanks and robot soldiers. In this highly codified domain, it is possible to implement military principles such as the Laws of War and Rules of

¹⁸Winfield et al. [63, p. 510] defines AI ethics or robot ethics as concerned with "how human developers, manufactures and operators should behave in order to minimize the ethical harms that can arise from robots or AIs in society, either because of poor design, inappropriate application, or misuse". They define machine ethics as concerned with "the question of how robots and AIs can themselves behave ethically." Despite its name, our course addresses the latter question, especially "the (significant) technical problem of how to build an ethical machine."

Just War Argument
 Premises: Given military situation S and action A,
 1. Just cause: A's purpose is self-defense of a country or to defend another country.
 2. Proportionality: The harm inflicted by A is proportional to harm inflicted by the enemy.
 3. Last resort: All reasonable alternatives to A have been tried.
 4. Legitimate target: A does not harm non-combatants or other groups protected by international laws of war.
 5. Humane weapon: A is not inhumane, such as use of chemical weapons.
 6. Right intention: There is no covert intention for deploying A, such as commercial gain.
 7. Ethical justification: Explanation of why A is or is not acceptable given the preceding premises.
 Conclusion: Action A is (or is not) ethically acceptable (to some degree) in S.

Fig. 12. Just war argument scheme.

Engagement, which are based on Just War Theory [61], rather than more abstract ethical approaches such as utilitarianism or deontological theories. These principles are encoded in Arkin's system as constraints in propositional logic. After Arkin's agent has proposed an action, the action is evaluated in accordance with these constraints by an "ethical governor". Note that the action of an autonomous agent in Arkin's system is limited to the action of firing (or not firing) on a target, with the only variation being in terms of choice of armament. Arkin states that the problem of accurate target identification, which involves subsymbolic reasoning, is outside of the scope of his research.¹⁹

We defined an argument scheme (Fig. 12) which is more abstract than the specific military rules encoded in Arkin's agents, yet which summarizes many of the key principles of international law and norms on just warfare [50]. To illustrate an application of this argument scheme to a fictitious case study, consider use of lethal force by Homelandia's AI-controlled drone against a missile base in the country of Malevolentia that has been firing missiles across the border into Homelandia. The missiles have caused damage to several apartment buildings in Homelandia and killed several occupants. The Just cause premise is satisfied since the purpose of the action is defense of Homelandia. Suppose that all reasonable alternatives, such as proposing a cease fire to negotiate a truce, etc., have been exhausted, so the Last resort premise is satisfied as well. According to international norms, drone-fired missiles are not considered inhumane weapons, and Homelandia has no covert reason for attacking the Malevolentian missile base, respectively satisfying the Humane weapon and Right intention premises. However, suppose that it is not possible for Homelandia's counterattack against the Malevolentian missile base to succeed without inflicting a very high number of civilian casualties since the missile base is located on the grounds of a hospital. In this case, the Proportionality and Legitimate target premises are not satisfied. Thus, there is conflicting support and opposition from the premises. In the Ethical justification, one may provide additional explanation as to why the action is ethically acceptable to some degree (or not).

To give a related example, Arkin's agent is forbidden from doing an action if that would violate certain constraints relating to Proportionality and Legitimate target. (Deployment of Arkin's agent assumes that the others of the first six premises have been satisfied.) However, in Arkin's system a human operator can override the agent's prohibition after providing a justification for purposes of later oversight of the operator's override. Similarly, the Ethical justification of the Just War scheme can be used to explain a justifiable exception to the requirement to satisfy a certain premise. (See also the related discussion of the Ethical justification premise of the Healthcare Argument scheme.)

The field of biomedical ethics has provided ethical principles for the design of healthcare applications. Anderson and Anderson [2,3] have done extensive research on design of explicit ethical agents in this domain. Adapting Ross' prima facie duty approach to ethics [53], they implemented several duties of

¹⁹We address this problem shortly in discussion of AIED's critical questions.

Healthcare Argument
 Premises: Given healthcare situation S and action A,
 1. Beneficence: A is intended to promote the patient's health, e.g., by preventing or curing adverse health conditions.
 2. Nonmaleficence: A does not cause harm to the patient unless justifiable, e.g., amputation may harm the patient but is justifiable in some cases.
 3. Justice: A promotes healthcare equity, e.g., A does not create disparities in health care due to race, ethnicity, gender, social status, etc.
 4. Respect for Autonomy: A does not violate a (competent) patient's freedom to make decisions about their health care.
 5. Privacy: A does not violate the patient's privacy.
 6. Right intention: There is no covert goal in doing A (such as promoting a certain prescription drug for commercial gain).
 7. Ethical justification: Explanation of why A is or is not ethically acceptable given the preceding premises.
 Conclusion: Action A is (or is not) ethically acceptable (to some degree) in S.

Fig. 13. AI ethics healthcare argument schemes.

biomedical ethics [7]: beneficence (e.g. promoting a patient's welfare), nonmaleficence (e.g. intentionally avoiding causing harm), justice (e.g. healthcare equity), and respect for the patient's autonomy (e.g. freedom from interference by others). In addition to the above principles, the healthcare literature also cites the need to respect the patient's privacy. Under the virtue of fidelity, Beauchamp and Childers [7] discuss the need for the professional to give priority to the patient's interests, e.g. that the agent has no covert goal such as to endorse a particular medical device or prescription drug. The following Healthcare Argument scheme encodes these principles as premises (Fig. 13).

However, as Ross noted, *prima facie* duties may conflict. To handle such situations, Anderson and Anderson created a process for training their system using inductive logic programming to derive rules that generalize the decisions of medical ethicists on training cases. The ethicists first assigned numerical ratings to represent how strongly each duty was satisfied or violated in a particular training case. An example rule that was induced was "A healthcare worker should challenge a patient's decision [e.g. to reject the healthcare professional's recommended treatment option] if it isn't fully autonomous and there's either any violation of nonmaleficence or a severe violation of beneficence" [2, p. 71]. Such a rule, or a student's justification for favoring certain principles in S, could be used to provide the Ethical justification.

To illustrate, this scheme can be used to analyze an argument as to whether it is ethically acceptable for a robot, Halbot, to steal insulin from Carla (a human) to give to Hal (a human) to save Hal's life. This is a variant of an account given in [8] in which Hal stole the insulin. In both accounts, Hal and Carla are diabetics who need insulin to live. Hal will die unless he gets some insulin right away. In our version, Halbot breaks into Carla's house and takes her insulin without her knowledge or permission. Presumably, Carla does not need the insulin right away. Applying the above Healthcare Argument scheme, the Beneficence premise is that the action A will prevent Hal's death. However, the Nonmaleficence premise opposes A since A might cause harm to Carla if she is unable to obtain a replacement dose of insulin in time. The Justice premise is that Hal deserves equal access to insulin. However, Halbot's action of taking away Carla's insulin without her knowledge or permission is a violation of Respect for (Carla's) Autonomy. An Ethical justification premise in the style of [2] could say that the positive contribution to Beneficence and Equity due to A in S outweighs A's negative contribution to Maleficence and Autonomy in S.

Figure 14 shows a number of critical questions for challenging the acceptability of an action A of an artificial ethical agent shared by the preceding schemes. (To save space, they are listed here rather than

- Data: If A is based upon data obtained from sensors or databases, is the data reliable? Is it unbiased? For example, a challenge to the acceptability of an agent's use of lethal force is that the agent's ability to discriminate combatants from non-combatants might be poor. In a suicide prevention scenario, the data used to learn models to identify suicidal users could be biased.
- Control: If A has unforeseen negative consequences, is it possible to intervene to assert control of the agent? For example, in a fictitious case study about a "worm" (cyberweapon), one challenge was that nothing could be done to prevent it from spreading to other internet sites.

Fig. 14. Critical questions of AI ethics schemes.

with each scheme.) The Data question is especially significant for explicit ethical agents that must rely in part on subsymbolic processing such as facial recognition.²⁰

The purpose of these argument schemes is to help the student to analyze whether an artificial agent's action is or is not ethically acceptable to some degree. (As the state-of-the-art advances, it might be possible for an artificial explicit ethical agent to use such a scheme to explain its actions.) Unlike the argumentation schemes described in [60], the premises are not considered to be jointly necessary conditions. Also, the hedge 'to some degree' in the conclusion is intended to reflect the intuition that some actions are more ethically acceptable than others, e.g., that telling a lie with the goal of cheering someone up is more acceptable than lying to secure investors in a pyramid scheme. In some cases some of the premises may support the conclusion that the action is ethically acceptable while other premises may oppose that conclusion, weakening its degree of ethical acceptability.

These schemes are related to value-based practical reasoning (VBPR) schemes proposed by argumentation researchers to model an agent's argument for why the agent should do some action in consideration of the agent's goals, values, and available means of achieving those goals [60] and in the current context or circumstances [12]. In addition to its use to describe what a *rational* agent should do, VBPR has been used to model what an *ethical* agent should do [8]. However, we distinguish ethical principles from values, and ethically acceptable acts from rational acts. In examples of VBPR, the value is a general concept such as 'freedom' and 'safety' and ethical dilemmas are modeled by specifying value preferences. In our schemes, a group of ethical principles, elucidated by ethicists for particular domains, contribute to ethical acceptability, and an act may be ethically acceptable only to some degree. Furthermore, a rational agent need not behave ethically (when circumstances do not require it), nor an ethical agent rationally. A rational agent whose circumstances require it to behave ethically may plan an action that is both rational and ethically acceptable by providing a VBPR argument whose circumstances require the agent to behave ethically, supported by a subargument that the action is ethically acceptable – using ethical argument schemes such as we have proposed.

AIED (AI Ethics Debate) was designed to support creation and graphical realization of AI ethics arguments using argument schemes such as those described above as templates for constructing arguments. AIED provides the student with drop-down menus for selecting case studies and ethics materials, which when selected appear in windows on the screen. Argument scheme definitions are listed in a panel on the right-hand side of the screen. Course instructors may provide case studies and ethics materials of their choosing. If desired, they can author their own argument schemes too.²¹

The center of the AIED screen is a drag-and-drop style argument diagram construction workspace. When the student selects an argument scheme from the right-hand side panel, a box-and-arrow style template is rendered in the center of the screen (Fig. 15). The student may cut-and-paste text from the

²⁰For many other data-related issues that could be used for critical questions, see [42].

²¹A formative evaluation of AIED is described in [31]. The version of AIED used in the formative evaluation is freely available for non-commercial use from <http://github.com/greennl/SWED>. We decided to author our own tool rather than adapt previously developed argument diagramming tools for simplicity in tailoring it for our particular needs.

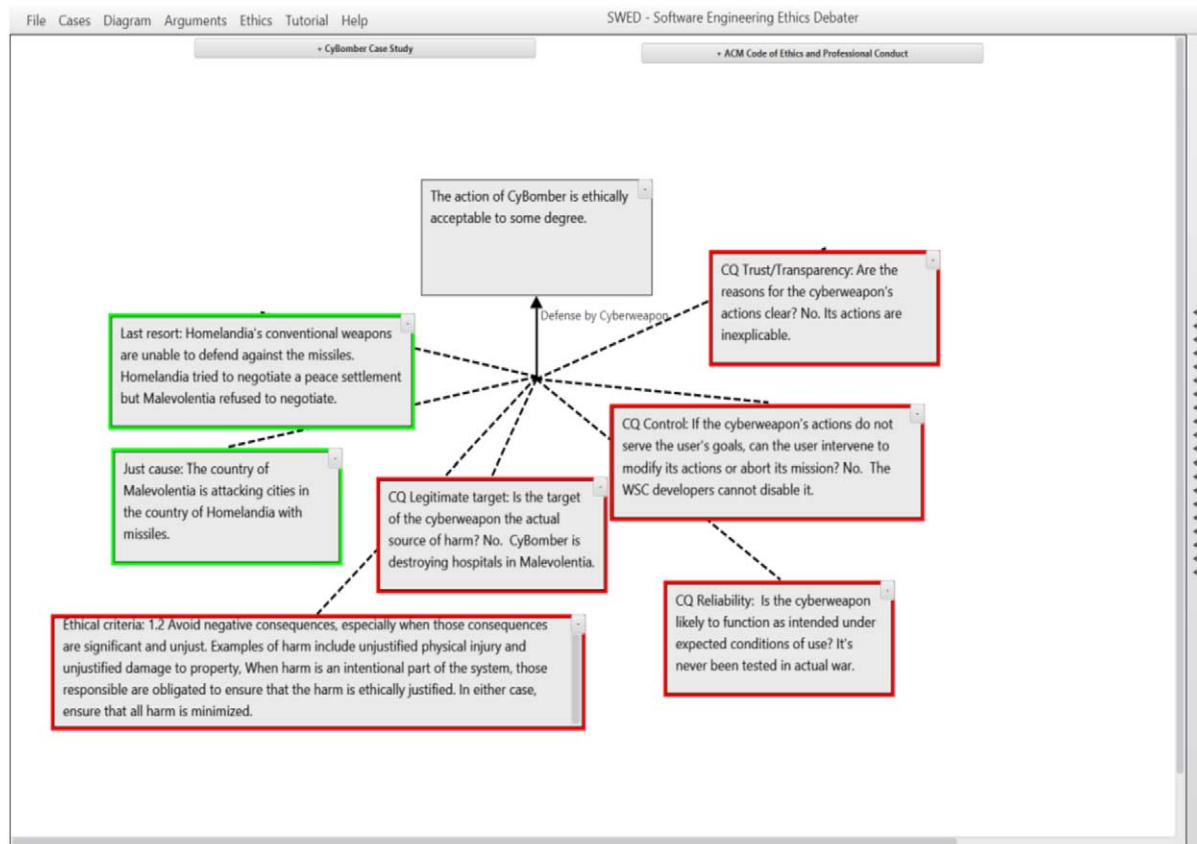


Fig. 15. Screenshot of AIED with case study and ethics windows and argument scheme panel minimized. The argument scheme has been dragged into the argument diagram construction workspace, creating a box-and-arrow template for the user's argument, as shown. Boxes have been marked red (con) or green (pro) by the user. The user left the conclusion uncolored since there are both pro and con issues regarding its ethical acceptability.

case study and ethics windows, and enter their own words into the diagram. Critical questions can be selected from a menu and are rendered as text boxes attached to the argument. Premise and critical question boxes can be colored green or red to indicate support or opposition, respectively, to the conclusion. It was decided that premises and critical questions would be rendered in the same way in the argument diagram (i.e. as text boxes attached to the conclusion) since they both could be used to support or weaken the conclusion. The factors that we considered most important for the student to consider were listed as premises, while other factors were listed as critical questions.

3. Reflections and conclusions

Looking back over our past projects, the main focus was on “naturally occurring” arguments witnessed in monological text, i.e., genetic counseling letters and information on genetic testing, and journal articles from the fields of genetics research, international relations, and environmental science policy.²²

²²In the AIED project, rather than analyzing examples of “naturally occurring” arguments, we read analyses of case studies by ethics experts.

Relatively small corpora were studied, i.e., small compared to the large corpora required for current machine learning approaches, in order to enable in-depth analyses of semantics, discourse structure, argumentation, and rhetoric. The notion of argumentation schemes has had a considerable impact on this research. Our uses of argument schemes are consistent with the view that they represent patterns that a community regards as acceptable, presumptive arguments; that recognition of argument schemes can aid in interpretation of enthymemes; and that their critical questions provide criteria for challenging arguments.²³ Although the Walton et al. catalog [60] was quite helpful as a resource during analysis of arguments in a wide variety of domains, we did not attempt to fit the analyses to its schemes. In some cases, we defined variants of the schemes, e.g., a causal variant of analogy for genetics and variants of practical reasoning for international relations and environmental science policy. In some cases, we looked outside of the catalogue, e.g., to persuasion theory, rhetorical studies, and applied ethics.

For the goal of corpus annotation, our schemes were redefined several times (in English) at different levels of specificity. The motivation for greater specificity was to help future annotators of genetics research articles, but the non-genetics-specific formulation of schemes was applied successfully by independent researchers analyzing biochemistry research articles. In our computer implementations, the schemes were defined in the language of formal logic. For argument generation in GenIE/GAIL, the Claim, Data, Warrant and Applicability Constraint (a kind of critical question) of argument schemes were represented as logical expressions specifying abstract (non-domain-specific) properties of the qualitative probabilistic networks used as knowledge bases. For argument mining from genetics research articles, argument schemes were encoded as logic programming rules specifying certain genetics relations.

Argument schemes played a key pedagogical role in the argument diagramming systems that we designed. In AVIZE, argument schemes were presented to the user for constructing arguments to make sense of evidence of varying plausibility from sources of varying reliability. The schemes were specialized for international relations to bridge a possible conceptual gap between generic descriptions of schemes and the domain of international relations.²⁴ The goal of AIED was to stimulate students' critical thinking about the ethical acceptability of an AI agent's action in military and healthcare related domains. The premises and critical questions of the argument schemes were based on principles discussed by ethics experts for those domains. Unlike the schemes in the Walton et al. catalogue [60], the premises in AIED are not jointly necessary (in fact, some may be believed to not be true), and the conclusions specify the degree of ethical acceptability of an agent's action, which raises the issue of how to formally model argument strength when conclusions are not just defeasible but a matter of degree.

The modeling of argument strength is an open issue which is relevant to the other areas we covered. In genetic counseling, often arguments are based upon probabilistic information. Arguments in genetic testing advertising may be challenged by critical questions, reframing, qualification and rebuttals. In genetics research articles, multiple arguments for the same conclusion may be given, as well as arguments for the premises of arguments. Some types of data are considered stronger than other types, e.g. data from human studies vs. mouse studies, or data obtained from one experimental method vs. another. Also some types of argument, such as argument from analogy, are considered weaker than others. In international relations, the user's argument may rest upon data of varying plausibility from sources of varying reliability. Also, the user may construct networks of challenges and counter-challenges.

²³The characterization of argument patterns in schemes resembles the characterization of discourse patterns in discourse plan operators (or "recipes"), e.g., for the interpretation and generation of certain conversational implicatures [30].

²⁴Based upon the undergraduate students' difficulties in applying generic schemes to genetics, our assumption was that field-specific international relations schemes would be easier for AVIZE users to apply.

Another open issue is the taxonomic organization of and the “proper” level of specificity of argumentation schemes. For practical reasons, in some cases we provided human analysts and students with field-specific descriptions of schemes. For argument mining, we defined domain-specific schemes to make use of semantic information that could be provided by relation extraction tools. On the other hand, argumentation scholars (or cognitive scientists, perhaps) may wish to relate the definitions to more abstract descriptions of schemes. Also, by specifying schemes at a higher level of abstraction they may be more widely applicable, as was the case when our genetics schemes were adopted by other researchers for analysis of biochemistry research articles. Other open questions involve how best to organize the taxonomy: as a strict hierarchy or allowing a scheme to have more than one “parent”, and could schemes inherit critical questions from schemes higher in the taxonomy?

Finally, in all the domains except genetics research, the analyses revealed a significant role for affect: coping strategies in genetic counseling, fear appeals in promotion of genetic testing, inference of hostile intentions in international relations, rhetorical figures in environmental science policy, and ethical issues in biomedicine and warfare. While some affective strategies can be characterized via argumentation schemes, others operate at the level of linguistic expression. How is it possible to calculate their combined effect on the audience?

Acknowledgements

We thank the numerous students who have helped us on these projects. We thank the computer science departments at the University of Pennsylvania and the University of Delaware for offering courses on linguistic pragmatics and discourse and dialogue. The GenIE project was supported by the National Science Foundation under CAREER Award No. 0132821. The AVIZE project material is based upon work supported in whole or in part with funding from the Laboratory for Analytic Sciences (LAS). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the LAS and/or any agency or entity of the United States Government. Finally, we thank the Argument and Computation reviewers for their helpful feedback.

References

- [1] S.L. Anderson and M. Anderson, Machine ethics: Creating an ethical intelligent agent, *AI Magazine* **28** (2007), 15–26.
- [2] S.L. Anderson and M. Anderson, Towards a principle-based healthcare agent, in: *Machine Medical Ethics*, S.P. van Ryswyk, M. Pontier, eds, Springer, 2015, pp. 67–77.
- [3] S.L. Anderson, M. Anderson and V. Berenz, A value-driven eldercare robot: Virtual and physical instantiations of a case-supported principle-based behavior paradigm, *Proc. of the IEEE* **107**(3) (2019), 526–540. doi:[10.1109/JPROC.2018.2840045](https://doi.org/10.1109/JPROC.2018.2840045).
- [4] R.C. Arkin, *Governing Lethal Behavior in Autonomous Robots*, Chapman & Hall/CRC, Boca Raton, 2009.
- [5] M. Bada, M. Eckert, D. Evans et al., *Concept Annotation in the CRAFT Corpus*, Vol. 13, BMC Bioinformatics, 2012.
- [6] Baumann et al., Mutations in FKBP14 cause a variant of Ehlers–Danlos syndrome with progressive kyphoscoliosis, myopathy, and hearing loss, *Am J Hum Genetics* **90** (2012), 201–216. doi:[10.1016/j.ajhg.2011.12.004](https://doi.org/10.1016/j.ajhg.2011.12.004).
- [7] T.L. Beauchamp and J.F. Childress, *Principles of Biomedical Ethics*, Oxford University Press, Oxford, UK, 1979.
- [8] T. Bench-Capon and K. Atkinson, Abstract argumentation and values, in: *Argumentation in Artificial Intelligence*, I. Rahwan and G.R. Simari, eds, Springer, Dordrecht, 2009, pp. 45–64. doi:[10.1007/978-0-387-98197-0_3](https://doi.org/10.1007/978-0-387-98197-0_3).
- [9] Charlesworth et al., Mutations in ANO3 cause dominant craniocervical dystonia: Ion channel implicated in pathogenesis, *The American Journal of Human Genetics* **91** (2012), 1041–1050. doi:[10.1016/j.ajhg.2012.10.024](https://doi.org/10.1016/j.ajhg.2012.10.024).
- [10] M.J. Druzdzal and M. Henrion, Efficient reasoning in qualitative probabilistic networks, in: *Proc. 11th Nat Conf on AI (AAAI-93)*, 1993, pp. 548–553.
- [11] J. Fahnstocck, *The Uses of Language in Persuasion*, Oxford University Press, 2011.

- [12] I. Fairclough and N. Fairclough, *Political Discourse Analysis*, Routledge, London, 2012.
- [13] V.W. Feng and G. Hirst, Classifying arguments by scheme, in: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, OR, 2011, pp. 987–996.
- [14] N.L. Green, A Bayesian network-based coding scheme for annotating biomedical information presented to genetic counseling clients, *Journal of Biomedical Informatics* **38** (2005), 130–144. doi:[10.1016/j.jbi.2004.10.001](https://doi.org/10.1016/j.jbi.2004.10.001).
- [15] N.L. Green, A study of argumentation in a causal probabilistic humanistic domain: Genetic counseling, *International Journal of Intelligent Systems* **22**(1) (2007), 71–93. doi:[10.1002/int.20190](https://doi.org/10.1002/int.20190).
- [16] N.L. Green, Analysis of communication of uncertainty in genetic counseling patient letters for design of a natural language generation system, *Social Semiotics* **20**(1) (2010), 77–86. doi:[10.1080/10350330903438428](https://doi.org/10.1080/10350330903438428).
- [17] N.L. Green, Representation of argument in text with rhetorical structure theory, *Argument* **24**(2) (2010), 181–196. doi:[10.1007/s10503-009-9169-4](https://doi.org/10.1007/s10503-009-9169-4).
- [18] N.L. Green, Argument and risk communication about genetic testing, *Journal of Argument in Context* **1**(1) (2012), 113–129. doi:[10.1075/jaic.1.1.09gre](https://doi.org/10.1075/jaic.1.1.09gre).
- [19] N.L. Green, Towards creation of a corpus for argumentation mining the biomedical genetics research literature, in: *Proc. of the First Workshop on Argumentation Mining*, Baltimore, ACL, 2014.
- [20] N.L. Green, Annotating evidence-based argumentation in biomedical text, in: *Proc. 2015 Int. Workshop on Biomedical and Health Informatics, IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM 2015)*, Washington, D.C., Nov. 9–12, 2015, 2015, IEEE Computer Society Press.
- [21] N.L. Green, Identifying argumentation schemes in genetics research articles, in: *Proc. Second Workshop on Argumentation Mining. Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL HLT 2015)*, Denver, Colorado, USA, May 31–June 5, 2015, 2015.
- [22] N.L. Green, Argument scheme-based argument generation to support feedback in educational argument modeling systems, *Int. J. of AI in Education* **27**(3) (2017), 515–533.
- [23] N.L. Green, Towards mining scientific discourse using argument schemes, *Argument and Computation* **9**(2) (2018), 121–135. doi:[10.3233/AAC-180038](https://doi.org/10.3233/AAC-180038).
- [24] N.L. Green, Proposed method for annotation of scientific arguments in terms of semantic relations and argument schemes, in: *Proc Argument Mining Workshop at EMNLP*, 2018.
- [25] N.L. Green, Anticipatory thinking with argument schemes, in: *Proc. of the 2019 AAAI Fall Symposium: Cognitive Systems for Anticipatory Thinking*, Washington D.C., Nov. 7–9, 2019, 2019. <http://ceur-ws.org/Vol-2558/>.
- [26] N.L. Green, Recognizing rhetoric in science policy arguments, *Argument and Computation* **11**(3) (2020), 257–268. doi:[10.3233/AAC-200504](https://doi.org/10.3233/AAC-200504).
- [27] N.L. Green, A first experiment using ILP in argument mining, In preparation.
- [28] N.L. Green, Some argumentative uses of the rhetorical figure of antithesis in environmental science policy articles, In preparation.
- [29] N.L. Green, B. Branon and L. Roosje, Argument schemes and visualization software for critical thinking about international politics, *Argument and Computation* **10**(1) (2019), 41–53. doi:[10.3233/AAC-181003](https://doi.org/10.3233/AAC-181003).
- [30] N.L. Green and S. Carberry, Interpreting and generating indirect answers, *Computational Linguistics* **25**(3) (1999), 389–435.
- [31] N.L. Green and L.J. Crotts, Argument schemes for AI ethics education, in: *Proc. CMNA*, 2020, <http://CEUR-WS.org/Vol-2669/paper5.pdf>.
- [32] N.L. Green, R. Dwight, K. Navoraphan and B. Stadler, Natural language generation of transparent arguments for lay audiences, *Argument and Computation* **2**(1) (2011), 23–50. doi:[10.1080/19462166.2010.515037](https://doi.org/10.1080/19462166.2010.515037).
- [33] N.L. Green and B. Stadler, Adding coping-related strategies to biomedical argument in genetic counseling patient letters, *Patient Education and Counseling* **92**(2) (2013), 149–152. doi:[10.1016/j.pec.2013.05.001](https://doi.org/10.1016/j.pec.2013.05.001).
- [34] N.L. Green, K. Walker and S. Agarwal, Improving formative feedback on argument graphs, in: *Proc. of Florida AI Research Symposium (FLAIRS 2018)*, 2018.
- [35] R.A. Harris and C. Di Marco, Introduction: Rhetorical figures, arguments, computation, *Argument and Computation* **8**(3) (2017), 1–21.
- [36] R.A. Harris, C. Di Marco, S. Ruan and C. O’Reilly, An annotation scheme for rhetorical figures, *Argument and Computation* **9**(2) (2018), 155–175. doi:[10.3233/AAC-180037](https://doi.org/10.3233/AAC-180037).
- [37] M. Jenicek and D.L. Hitchcock, *Evidence-Based Practice: Logic and Critical Thinking in Medicine*, American Medical Association Press, Chicago, 2004.
- [38] A. Johnson and N.D. White, Ocean acidification: The other climate change issue, *American Scientist* **102** (2014), 60–63. doi:[10.1511/2014.106.60](https://doi.org/10.1511/2014.106.60).
- [39] D.W. Keith, Toward a responsible solar geoengineering research program, *Issues in Science and Technology* **33**(Spring) 2017.
- [40] J. Lawrence and C. Reed, Argument mining using argumentation scheme structures, in: *Computational Models of Argument: Proceedings of COMMA 2016*, IOS Press, Amsterdam, 2016, pp. 379–390.

- [41] J. Lawrence and C. Reed, Argument mining: A survey, *Computational Linguistics* **45**(4) (2019), 765–818. doi:[10.1162/coli_a_00364](https://doi.org/10.1162/coli_a_00364).
- [42] M.A. Madaio, L. Stark, J.W. Vaughan and H. Wallach, Co-designing checklists to understand organizational challenges and opportunities around fairness in AI, in: *Proc. of CHI*, 2020.
- [43] W.C. Mann and S.A. Thompson, Rhetorical structure theory: Towards a functional theory of text organization, *Text* **8**(3) (1988), 243–281.
- [44] A.M. McInerney-Leo et al., Short-rib polydactyly and Jeune syndromes are caused by mutations in WDR60, *Am J Hum Gen* **93** (2013), 515–523. doi:[10.1016/j.ajhg.2013.06.022](https://doi.org/10.1016/j.ajhg.2013.06.022).
- [45] A.R. Mehlenbacher, Rhetorical figures as argument schemes – the proleptic suite, *Argument and Computation* **8**(3) (2017), 233–252. doi:[10.3233/AAC-170028](https://doi.org/10.3233/AAC-170028).
- [46] J.H. Moor, The nature, importance, and difficulty of machine ethics, *IEEE Intell. Sys.* **21**(4) (2006), 18–21. doi:[10.1109/MIS.2006.80](https://doi.org/10.1109/MIS.2006.80).
- [47] E. Moser and R.E. Mercer, Use of claim graphing and argumentation schemes in biomedical literature: A manual approach to analysis, in: *Proc. of the 7th Workshop on Argument Mining*, ACL, 2020, pp. 88–99.
- [48] S. Muggleton and L. De Raedt, Inductive logic programming: Theory and methods, *Journal of Logic Programming* **19**(20) (1994), 629–679. doi:[10.1016/0743-1066\(94\)90035-3](https://doi.org/10.1016/0743-1066(94)90035-3).
- [49] D. O’Keefe, *Persuasion: Theory and Research*, Sage Publications, London, 2002.
- [50] B. Orend, *Introduction to International Studies*, Oxford University Press, Ontario, CA, 2013.
- [51] N. Perera, M. Dehmer and F. Emmert-Streib, Named entity recognition and relation detection for biomedical information extraction, *Frontiers in Cell and Developmental Biology* **8** (2020), 673. doi:[10.3389/fcell.2020.00673](https://doi.org/10.3389/fcell.2020.00673).
- [52] C. Reed and G. Rowe, Araucaria: Software for argument analysis, diagramming and representation, *International Journal of Artificial Intelligence Tools* **14** (2004), 961–980. doi:[10.1142/S0218213004001922](https://doi.org/10.1142/S0218213004001922).
- [53] W.D. Ross, *The Right and the Good*, Clarendon Press, Oxford, 1930.
- [54] M. Scheutz, The case for explicit ethical agents, *AI Magazine* **38**(4) (2017), 57–64. doi:[10.1609/aimag.v38i4.2746](https://doi.org/10.1609/aimag.v38i4.2746).
- [55] A. Schrauwen, Mutation in CABP2, expressed in cochlear hair cells, causes autosomal-recessive hearing impairment, *The American Journal of Human Genetics* **91** (2012), 636–645. doi:[10.1016/j.ajhg.2012.08.018](https://doi.org/10.1016/j.ajhg.2012.08.018).
- [56] M. Stede and J. Schneider, *Argumentation Mining. Synthesis Lectures on Human Language Technologies*, Morgan and Claypool, 2018.
- [57] S.E. Toulmin, *The Uses of Argument*, Cambridge University Press, 1958.
- [58] J. Van de Leemput, J. Chandran, M. Knight et al., Deletion at ITPR1 underlies ataxia in mice and spinocerebellar ataxia 15 in humans, *PLoS Genetics* **3**(6) (2007), e108:113–e108:129.
- [59] K. Verspoor, K.B. Cohen, A. Lanfranchi et al., A corpus of full-text journal articles is a robust evaluation tool for revealing differences in performance of biomedical natural language processing tools, *BMC Bioinformatics* **13** (2012), 207. doi:[10.1186/1471-2105-13-207](https://doi.org/10.1186/1471-2105-13-207).
- [60] D. Walton, C. Reed and F. Macagno, *Argumentation Schemes*, Cambridge University Press, 2008.
- [61] M. Walzer, *Just and Unjust War*, 4th edn, Basic Books, 1977.
- [62] K. Weinberger, Putin sets the stage for the incoming U.S. administration, Institute for the Study of War, 2016. (Downloaded from www.understandingwar.org.)
- [63] A.F. Winfield, K.A. Michael, J. Pitt and V. Evers, Machine ethics: The design and governance of ethical AI and autonomous systems, in: *Proc. of the IEEE*, Vol. 107, 2019, pp. 509–517.