Academic Writing Support System Using Bayesian Networks

Conference Paper · July 2015
DOI: 10.1109/ICALT.2015.16

2 authors:

Masaki Uto
The Univ. of Electro-Communications
15 PUBLICATIONS **33** CITATIONS
SEE PROFILE

Maomi Ueno
The University of Electro-Communications
121 PUBLICATIONS **372** CITATIONS
SEE PROFILE

Academic Writing Support System using Bayesian Networks

Masaki Uto
Nagaoka University of Technology
Niigata, Japan
uto@oberon.nagaokaut.ac.jp

Maomi Ueno
University of Electro-Communications
Tokyo, Japan
ueno@ai.is.uec.ac.jp

***Abstract*—For academic writing, elaborating an argument particularly addressing an *argument strength* is important to establish causal relations between sentences. However, when an argument becomes large or complex, elaborating an argument considering the *argument strength* is difficult. To solve this problem, this article presents a proposal for an argument elaboration support system using a Bayesian network representation of the Toulmin model. Using that Bayesian network representation, the proposed system can estimate *argument strength*, *sentence validity*, and *sentence influence*. Moreover, it can generate optimal advice for revising the argument.**

***Keywords*—Toulmin model; Bayesian network; academic writing; argument; information system**

I. INTRODUCTION

Academic writing is intended to convey information effectively [1]. To accomplish that aim, constructing a persuasive argument is necessary. Constructing a persuasive argument requires the establishment of a causal relation between sentences and necessitates elaboration of the argument objectively and repetitively. Nevertheless, these activities are known to be difficult, especially for beginners.

For this reason, various argument elaboration support systems have been developed for academic writing [2]. Most of these systems support the construction of an argument fitting an argument scheme called the Toulmin model[3]. The Toulmin model decomposes an argument into six components: *claim*, *data*, *warrant*, *backing*, *qualifier*, and *rebuttal*. It also formalizes the relation between these components as a directed graph. The Toulmin model has been used in various fields as a standard for evaluating arguments. Its validity has been demonstrated.

Although previous systems have remained limited to support the superficial fitting of an argument to the Toulmin model, it is more important for construction of a persuasive argument to elaborate an argument by particularly addressing the *argument strength*, which is the strength of the causal relation between sentences. However, when an argument becomes large and complex, the following elaboration activities corresponding to the *argument strength* become difficult. 1) Evaluating argument strength among all sentences. 2) Estimating the validity of all sentences. 3) Determining how each sentence influences the justification of the claim.

To resolve these problems, this article presents development of an argument elaboration support system using a Bayesian network (*BN*) representation of the Toulmin model. This system expresses sentences in the argument as probabilistic variables and the causal relations between sentences as conditional probabilities. Then, an argument constructed to fit to the Toulmin model can be formalized as a BN. Probability is expressed as subjective probability, which users evaluate using six levels of categories corresponding to numerical probability values. Using the BN representation, the proposed system can realize the following features. 1) For all arguments and sentences, three argumentation characteristic indexes, namely, *argument strength*, *sentence validity*, and *sentence influence on the claim*, can be estimated. 2) Feedbacks on revising the argument can be provided in accordance with the estimated argumentation characteristic indexes

This article also describes the effectiveness of the proposed system through results of subjective experiments.

II. TOULMIN MODEL

For this study, we use the Toulmin model[3] as a normative model of argumentation. The Toulmin model decomposes an argument into six components: *claim*, *data*, *warrant*, *backing*, *qualifier*, and *rebuttal*. Then it formalizes the relation between these components as a directed graph as presented in Fig. 1. Here, *claim* is the position being argued for. *Data* represents evidence used to support the claim. *Warrant* stands for hypothetical reasoning that provides a bridge between data and claim. *Backing* is a statement that gives authority and credibility to the warrant. *Qualifier* expresses the strength of justifying the movement from data to claim. *Rebuttal* explains conditions in which the general validity of the warrant does not hold.

The Toulmin model T is defined as $T =< \boldsymbol{N},\ \boldsymbol{E} >$, where set $\boldsymbol{N}$ consists of the components of the Toulmin model and directed edge set $\boldsymbol{E}$ expresses the supporting relations among the components[4]. Here, the directed edges cannot be assigned arbitrarily. They must follow fixed rules, or *formation rules*. To define formation rules, we extracted all rules from previous studies(e.g., [4], [5]). The extracted rules are presented in Fig. 2.

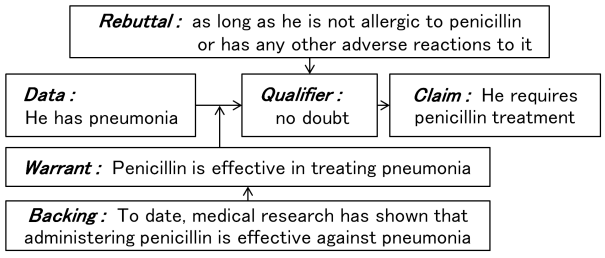

Figure 1. Toulmin model.

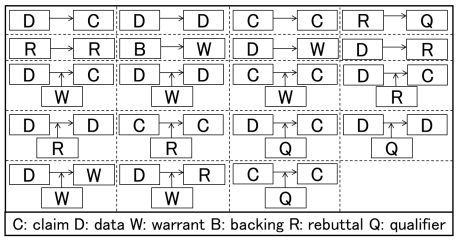

Figure 2. Toulmin model formation rules.

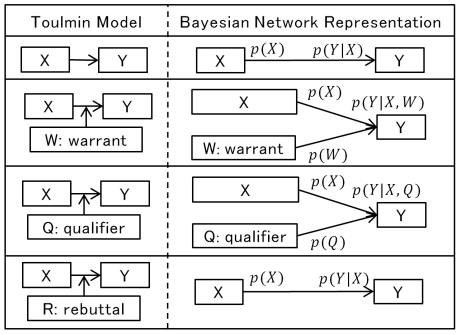

Figure 3. BN representation of Toulmin model.

In this study, a Toulmin model constructed following the formation rules is represented as the BN.

III. Bayesian network representation of Toulmin model

The persuasiveness of a sentence within an argument generally cannot be judged as true or false. The persuasiveness must be evaluated using a probabilistic criterion indicating the degree of propriety in the sentence (referred to below as *validity*). Accordingly, this study regards each sentence within an argument as a probabilistic variable S_i and expresses the validity of a sentence as subjective probability $p(S_i = k)$. Here, $k \in \{0,1\}$ expresses the state of a variable, where 1 means true and 0 means false. In addition, the validity of sentence S_i given sentence set $\mathbf{\Pi}_{S_i}$ is expressed as conditional probability $p(S_i = k|\mathbf{\Pi}_{S_i} = j)$. Here, $\mathbf{\Pi}_{S_i} = j$ shows that $\mathbf{\Pi}_{S_i}$ takes on the j-th pattern.

The BN represents the relations among probabilistic variables. The BN is defined as $B =< \mathcal{G}, \boldsymbol{p} >$. Here, $\mathcal{G} =< \boldsymbol{V}, \boldsymbol{E} >$ is a directed acyclic graph consisting of vertex set $\boldsymbol{V}$ corresponding to a variable set and directed edge set $\boldsymbol{E}$ expressing the probabilistic dependency among those variables. $\boldsymbol{p}$ expresses the conditional probability distributions corresponding to each variable.

In this study, the BN structure $\mathcal{G}$ is determined from a Toulmin model constructed by a user. The transformation from the Toulmin model to a BN structure is conducted to reflect the logical relations between the Toulmin model components. This study defines the transformation rules as presented in Fig. 3.

The conditional probabilities $\boldsymbol{p}$ are assessed by users using words corresponding to subjective probability values. The words and the probability values are defined as {sufficiently valid (0.95), probably valid (0.77), if anything, valid (0.59), if anything, not valid (0.41), not very valid (0.23), or not valid at all (0.05)} (hereinafter, *assessment categories*).

IV. Argumentation Characteristic Indexes

This section presents the three argumentation characteristic indexes using the BN representation.

***Argument strength*:** Because the *argument strength* refers to the strength of the causal relation between sentences, this study formalizes it using conditional mutual information (*MI*) that expresses the strength of the dependency between variables. The MI takes on a larger value when the dependency among variables is strong. This study judges an argument with the MI < 0.05 as weak.

***Sentence validity*:** Because the *sentence validity* represents the degree of justification for a sentence in the argument, this study defines it as the marginal probability of the variable being true. In this study, a sentence is regarded as invalid when its validity < 0.7.

***Sentence influence*:** *Sentence influence* indicates how a sentence S_i influences the validity of claim S_c. More specifically, it expresses how the validity of the claim S_c will change when the validity of sentence S_i changes. A method to realize such analysis in the BN is known as the *sensitivity analysis*[6]. In the sensitivity analysis, $p(S_c = 1)$ is calculable by a function of $p(S_i = 1)$, called *sensitivity function*. The sensitivity functions enable users to analyze the *sentence influence*.

V. Argument elaboration support system

This section presents a description of the system that was ultimately developed. The system interface is depicted in Fig. 4. A user inputs a text using any text editor (e.g., upper left window in Fig. 4) and uses this system (right window in Fig. 4) to construct and elaborate an argument. According to instructions displayed in the right part of the system, the user input sentences corresponding to each Toulmin component to the system. Then, the Toulmin model is visualized in the left part of the system. In the process of constructing a Toulmin model, the user is asked to self-assess the validity of sentences using a self-assessment window like the window below left in Fig. 4. After the self-assessment, the system calculates the argumentation characteristic indexes for all nodes and highlights any nodes and edges corresponding to a low validity sentence and a weak argument. If the user clicks a node or edge, then the argumentation characteristic indexes and advice for revising the argument are displayed in the right part of the system. Here, the advice is generated in accordance with the estimated argumentation characteristic indexes by following the rules presented in Table I. The user elaborates the text and argument by reference to the argumentation characteristic indexes and advice.

VI. System Evaluation Experiment

This section evaluates whether the system can support the elaboration of a large argument considering *argument strength*, *sentence validity*, and *sentence influence* through the following subjective experiments.

In this experiment, 20 university students were collected as subjects. The subjects were divided randomly into group A using the proposed system and group B using a system that only had a function of visualizing the Toulmin model. Each group consisted of 10 subjects. After explanations related to

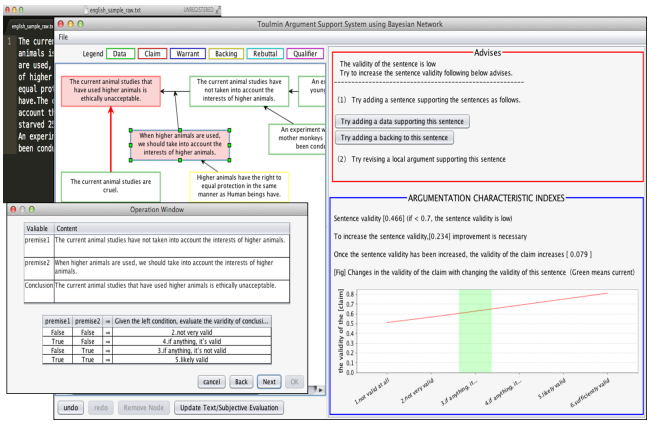

Figure 4. Toulmin model formation rules.

TABLE I. ADVICE GENERATION RULES.

Advice for a sentence with low validity

- The validity of the sentence is low.
- If the sentence validity is increased,
  then the validity of the claim increases $\triangle p(S_c = 1) = (\mu_i \varepsilon_1 + \nu_i) - p(S_c = 1)$.
- To increase the sentence validity, $\varepsilon_1 - p(S_i = 1)$ improvement is necessary.
- Try to increase the sentence validity following the advice shown below.
  - · Try adding a sentence supporting the sentences as follows.
    - + Try adding data supporting this sentence
    - + Try adding a warrant to this argument
  - · Try revising a local argument supporting this sentence

Advice for an argument with weak strength

- The argument is weak. Try to improve the argument following the advice below.
  - · Remove the argument and try adding a stronger argument.
  - · Try adding a sentence supporting the argument as follows.
    - + Try adding a warrant to this argument
    - + Try adding a qualifier to this argument

the Toulmin model, argumentation characteristic indexes and system operation, we provided the hard copies of a relatively large argument and of its Toulmin model. Each subject was asked to construct the same structure of the Toulmin model shown in the hard copy using the systems. Then, the subject was asked to perform the following activities using the system freely. (1) Show all sentences and arguments in which revisions are regarded as necessary (designated as *designated locations*). (2) For all *designated locations*, note the reason why a revision is probably necessary (designated as *designated reason*) and how it should be revised (designated as *revision method*). (3) For all sentences in the argument, evaluate the importance of the sentence for the claim (designated as *influence*) using three scales (1. important, 2. fair, 3. unimportant). (4) Revise the arguments.

To evaluate the validity of *designated reason*, *revision method*, and *influence* made by the subjects, the responses of the subjects were compared to those of an expert. To obtain expert responses, an expert asked to perform the same activities under group B's condition. Here, we categorized the freely described *designated reasons* and *revision methods*. Then, between the responses of each subject and the expert, the κ coefficient, which is an index expressing the rate of agreement between judgments made by multiple evaluators, was calculated.

The average values and standard deviations of the calculated κ coefficients were presented in Table II. In Table II, $p < .05(.01)$ expresses that a significant difference

TABLE II. EXPERIMENTAL RESULTS.

	κ coefficient			Expert evaluation	
	Designated reason	**Revision method**	**Influence**	**Argument strength**	**Sentence validity**
group A	.39(.15)	.36(.18)	.38(.22)	3.70(.79)	3.70(.79)
group B	.21(.18)	.18(.17)	.15(.13)	2.10(1.10)	2.10(1.10)
	$p < .05$	$p < .05$	$p < .01$	$p < .01$	$p < .01$

occurred at a level of significance of 5% (1%). According to the results, the κ coefficients for group A were significantly larger than group B. Therefore, the proposed system enables subjects to elaborate an argument similar to that of an expert in terms of the argument characteristics.

Furthermore, to evaluate whether the revised arguments were improved in terms of *argument strength* and *sentence validity*, an expert evaluated the revised argument in the following two items using a four-level scale (1. numerous, 2. many, 3. not many, 4. almost none). (1) Are there any arguments with weak argument strength that require revision? (2) Are there any sentences with low validity that require revision?

The results are listed in Table II. These results show that group A scored significantly higher, meaning that the system enables subjects to revise arguments in terms of *argument strength* and *sentence validity*.

VII. CONCLUSION

This article proposed a system that supports the elaboration of an argument particularly addressing the *argument strength*. To realize this feature, this article proposed BN representation of the Toulmin model and the three argumentation characteristic indexes. Additionally this article proposed a method to provide advice on revising an argument based on the argumentation characteristic indexes. Through the subjective experiments, this article demonstrated that the proposed system can support the elaboration of an argument addressing the *argument strength*.

This article does not assert that the proposed system can promote to acquire the argument elaboration skill. In future research, we plan to extend the system to promote the acquisition of the skill.

REFERENCES

[1] N. Kando, "Text-level structure of research papers: Implications for text-based information processing systems," in *In Proceedings of the British Computer Society Annual Colloquium of Information Retrieval Research*, 1997, pp. 68–81.

[2] O. Scheuer, F. Loll, N. Pinkwart, and B. M. McLaren, "Computer-supported argumentation: A review of the state of the art," *International Journal of Computer-Supported Collaborative Learning*, vol. 5, no. 1, pp. 43–102, 2010.

[3] S. E. Toulmin, *The Use of Argument*. Cambridge University Press, 1958.

[4] C. Reed and G. Rowe, "Araucaria: Software for argument analysis, diagramming and representation," *International Journal of AI Tools*, vol. 14, pp. 961–980, 2004.

[5] B. Verheij, "Evaluating arguments based on toulmin's scheme," *Argumentation*, vol. 19, no. 3, pp. 347–371, 2005.

[6] H. Chan and A. Darwiche, "Sensitivity analysis in bayesian networks: from single to multiple parameters," in *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, ser. UAI '04, 2004, pp. 67–75.