

Determination of Reflective User Engagement in Argumentative Dialogue Systems

Annalena Aicher¹, Wolfgang Minker¹ and Stefan Ultes²

¹Ulm University, Institute of Communications Engineering, Albert-Einstein-Allee 43, 89081 Ulm, Germany

²Mercedes Benz AG, Stuttgart, Germany

Abstract

In this work we propose to our knowledge the first approach to determine the reflective user engagement (RUE) during an argumentative dialogue. Therefore, we review state-of-the-art literature definitions for reflective engagement (RE) and approaches to measure the latter. Given some basic characteristics the argumentative dialogue system has to provide, we derive a formula to determine the RE taking into account the argument structure and the respective current position at each state of the dialogue.

Keywords

Reflective User Engagement, Argumentative Dialogue Systems, Bipolar Argumentation Structures

1. Introduction

A natural way of resolving different points of view or forming an opinion for humans is through conversation, i.e., through the exchange of arguments. Due to the vast amount of different available information people tend to focus on a biased subset of sources that repeat or strengthen an already established or convenient opinion which is furthermore reinforced by filter algorithms [20].

In order to avoid the (often unconscious) process of intellectual isolation, we suggested an approach to explore large amounts of diverging information in a natural and intuitive way [1]. On this basis we aim for a system that provides an engaging form of interaction via natural language and encourages users to address diverging points of view and to scrutinize information. In order to foster a dialogue conveying a balanced discussion of topics, we will extract reward signals required for reinforcement learning from properties of the argumentative dialogue between the user and the system. In particular one property is the RUE, denoting the critical-thinking and open-mindedness demonstrated by the user in the interaction with the system. In their study [16] Masrek et al. showed that user engagement is a strong predictor of user satisfaction and thus, crucial to keep the users motivated to talk to the system and confront themselves with diverging arguments. Therefore, we derive an in-dialogue calculation for RUE taking into account the argument structure and user behavior during the dialogue.

The remainder of this paper is as follows: the overview over the related work in Section 2

CMNA'21: Workshop on Computational Models of Natural Argument, September 2-3, 2021, Online

✉ annalena.aicher@uni-ulm.de (A. Aicher); wolfgang.minker@uni-ulm.de (W. Minker); stefan.ultes@daimler.com (S. Ultes)

ORCID 0000-0002-5634-5556 (A. Aicher); 0000-0003-4531-0662 (W. Minker); 0000-0003-2667-3126 (S. Ultes)

 © 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

is followed by Section 3 describing our proposed derivation of the RUE after explaining the dialogue model upon which the former is based. In Section 4 we conclude by summarizing the presented ideas and give a short outlook.

2. Related Work

In general, O'Brien et al. [18] define engagement as the 'quality of user experiences with technology that is characterized by challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control and time, awareness, motivation, interest, and affect'. Lalmas et al. [13] specify user engagement to be the quality of the user experience that emphasizes the positive aspects of interacting with an online application and the desire to use it longer and repeatedly.

As user engagement is a very complex phenomenon there exist numerous of (potential) measurement approaches. Common ways to evaluate user engagement include using self-report measures like questionnaires [7, 21]; observational methods such as facial expression analysis [6]; neuro-physiological signal processing methods for example cardiovascular accelerations [13]. Oh et al. [19] suggest a measurement and structural model for empirically capturing the meaning and process of user engagement in the context of interactive media. They chose four attributes, i.e. physical interaction, interface assessment, absorption, and digital outreach.

Other studies ([2, 4]) examined the variety of time measures, cursor movement, and eye tracking data, in addition to self-reported items and click data. Lalmas et al. [13] give an overview on techniques based on physiological measurement, such as bodily and brain response and function [17] and eye tracking [3]. Correlations between gaze tracking and cursor tracking are discussed by [10]. Measures based on web analytics include online behavioral metrics e.g. click-through rates [22], number of page views [11], time spent on a site (i.e., dwelltime [8, 26]) and frequency of return visits [14].

Still, as stated by Arapakis et al. [2] it is important to move beyond the 'legacy of the click' and consider cognitive and affective factors of engagement. Silpasuwanchai [25] et al. relate cognitive engagement to the sense of involvement, focused attention, and deep reflection.

Prado-Romero et al. [23] propose to use anomaly detection for finding 'influential' and 'open minded' individuals in the Twitter network. Their approach is based on the InterScore anomaly detection algorithm, identifying users with an anomalous number of out- and in-edges. According to Haim et al. [9] open-mindedness correlates with linguistic style accommodation¹ and relates to the assumed speaker role in different contexts. In contrast to our work, these approaches are not (preliminary) concerned with determining content-related open-mindedness. According to [5, 15, 24] reflective engagement (RE) refers to learners' continual and active participation in their problem inquiry with a continuous and critical judgment of inquiry process and inquiry outcomes for possible improvement. Most approaches that describe RE are strongly connected to teaching-learning processes [12, 25]. Instead we consider a more general definition, which refers to the user's motivation in scrutinizing arguments and exploring diverging views. In extension to existing literature we propose a calculation approach extracted from the user

¹Linguistic style accommodation denotes the 'unconscious process in which a speaker accommodates their communicative behavior with respect to the communication partner' [9]

behavior (actions) instead of solely relying on self-report measures.

3. Reflective User Engagement in BEA

In the following we shortly point out the main characteristics of the dialogue model used in the argumentative dialogue system BEA [1]. Based on this model we then calculate the RUE.

3.1. Dialogue model

The interaction between the ADS and the user is separated in turns, consisting of a user action and corresponding natural language answer of the system. The possible actions (moves) the user is able to choose from, depend on the position of the current argument (root / parent node / 'leaf' node). Due to limited space we will focus only on the moves which are relevant to derive the RUE.

To prevent the user from being overwhelmed by the amount of information, the user is able to navigate incrementally through the argument structure resembling the one of a tree based on bipolar argument structures. These structures depict *support* or *attack* relations between the arguments (nodes) in a graph. We choose a non-cyclic tree structure, where each node ('parent') is supported or attacked by its 'children'. If no children exist, the node is a leaf and marks the end of a branch. Usually a single *major claim* formulates the overall topic, representing the root node in the graph.

The user is able to specify if he enquires for a supporting (pro) or attacking (con) argument on the current argument. For a better understanding, we will consider the following example. Let the topic of the discussion be concerned with the question whether to stay in a certain hotel or not. One aspect of the discussion might be the service of the hotel. Thus, the user can e.g. request more information by stating: 'I would like to hear a supporting/contradicting argument for the claim, that the service of the hotel is very good.'

At any time during the conversation the user is able to ascend the argument branch (level up to the 'parent' node) and descend on another unknown branch (targeting the parent node) again. But in doing one will not be able to return to the previous branch, especially if one has not heard all arguments, these arguments will be 'dropped'. In this case we assumed that either the user lost interest in the current argument or received in his/her perception sufficient information. This is important to keep in mind for the following derivation.

3.2. Derivation of the Reflective User Engagement

We propose an approach based on Yi et al. [26], who correlate rather short website content and long browsing time with great user interest. In analogy to this a user who inquires for more information is more engaged. Recalling our previous definition of reflective engagement as the user's interest scrutinizing arguments and exploring diverging views. This can be mapped to the two actions of the user asking for more information, either pro or con sides of the current argument². Thus, the more arguments of both sides are heard, the higher is the RUE. The

²BEA visualizes all subtrees of the current argument, such that the user knows exactly how many arguments are available. This is crucial as we assume that unvisited arguments are intended and not just missed by mistake

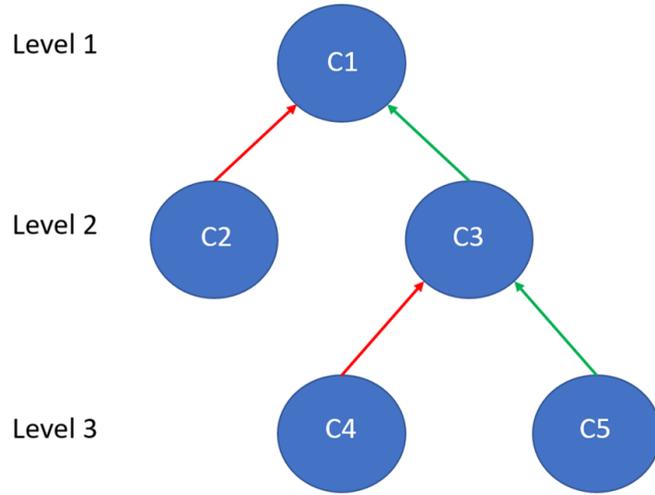


Figure 1: Argument subtree structure consisting of five Claims (C1-C5), with supporting (denoted in green) and attacking (denoted in red) relations.

highest RUE is given if the same number of pro and con arguments are heard. To take a potential, data-related bias ($\#pro \neq \#con$) into account, we introduce the characteristic function $\mathbb{1}_{p \text{ visited}}$. It considers if at least one pro/con pair has been heard and if so, makes it possible to consider single additional arguments, which have been heard. Thus, we define:

$$\mathbb{1}_{p \text{ visited}} = \begin{cases} 1, & \text{if } \exists \text{ visited pro/con pairs} \\ 1, & \text{if no pro/con pairs exist} \\ 0, & \text{if } \nexists \text{ visited pro/con pairs} \end{cases} \quad (1)$$

For example, if we consider the simple argument subtree structure shown in 1, on each level both arguments (pro and con) have to be heard such that the characteristic function $\mathbb{1}_{p \text{ visited}} = 1$. If only one side is heard, e.g. solely C2 or only C3 it follows $\mathbb{1}_{p \text{ visited}} = 0$ for the respective Level 2. Likewise this follows for Level 3, in case just C4 or C5 are heard.

As the RUE reflects critical thinking and openmindedness of the user, we weight a balanced relation of pro and con pairs higher than the exploration of solely the pro or con side of an argument. We choose to weight all visited pro/con pairs with a factor $\alpha_k > 0.5$ and all single arguments with $(1 - \alpha_k) < 0.5$. Without loss of generality, if no pro/con pairs exist for level $k + 1$ it follows:

$$\alpha_k := 0; \quad \frac{\#P^{(k+1)\text{visited}}}{\#P^{(k+1)\text{all}}} := 0$$

and vice versa, if no single arguments exist

$$\alpha_k := 1; \quad \frac{\#S^{(k+1)\text{visited}}}{\#S^{(k+1)\text{all}}} := 0.$$

The α is recommended to be chosen depending on the relation between pro/con pairs and single pro or con arguments.

The resulting RUE of a parent node r_k for the single level k can therefore be determined by:

$$r_k = \alpha_k \frac{\#p^{(k+1)}\text{visited}}{\#p^{(k+1)}\text{all}} + \mathbb{1}_{\text{p visited}} (1 - \alpha_k) \frac{\#s^{(k+1)}\text{visited}}{\#s^{(k+1)}\text{all}}, \quad (2)$$

where $\#p^{(k+1)}$ denotes the number of child pro/con pairs at level $k + 1$ and $s^{(k+1)}$ denotes the number of single children at level $k + 1$. Regarding the given example in Figure 1, for $k = 1$ it follows, that $r_1 = \alpha_k \frac{1}{1} = 1$ if both C2 and C3 are heard. If only C2 or C3 are heard $r_1 = 0$. For r_2 can be derived completely analogously.

When considering hierarchical argumentation structures, arguments at the beginning of a branch are more general than ones at deeper levels. Due to this we introduce a hierarchical weight ω_d in order to incorporate the different levels of argument depth into our reflective engagement measure.

Therefore, a balanced exploring of lower levels will be assigned larger weight values than near the root node.

$$\omega_{d,k} = \frac{d_{l,(k-j)}}{\sum_{m=1}^{l_{\max}-j} m}, \quad (3)$$

where $d_{l,(k-j)}$ denotes the depth of the level k with respect to the level of parent node at level j . If we look e.g. at C3 with $k = 1$ and $j = 2$ and assume all C1-C5 have been heard, we get

$$\omega_{d,1} = \frac{1}{\sum_{m=1}^{3-1} m} = \frac{1}{3}.$$

To avoid an over-representation of levels with only few arguments while levels with many arguments will be under-represented, we define a weight ω_n which takes the different sizes of levels into account. Thus, we relate the number of descendants of the respective level k to all descendants such that

$$\omega_{n,k} = \frac{\#\text{pro}_k + \#\text{con}_k}{\sum_{m=k+1}^{l_{\max}} \#\text{pro}_m + \#\text{con}_m}, \quad (4)$$

where $\#\text{pro}_k, \#\text{con}_k$ denotes the number of all pro, con arguments. Again assuming that all claims C1-C5 have been heard, it follows for the level $k = 1$ that $\omega_{n,1} = \frac{2}{4} = 0.5$. For the overall RUE at the parent node at j it follows:

$$RUE_j = \frac{\sum_{j=k}^{l_{\max}-1} \omega_{d,k+1} \omega_{n,k+1} r_k}{\sum_{j=k}^{l_{\max}-1} \omega_{d,k+1} \omega_{n,k+1}}, \quad RUE_j \in [0, 1], \quad (5)$$

which denotes the normalized sum over the weighted reflective user engagement values r_j for each descending level $j + 1, j + 2, \dots, l_{\max}-1$ ³. Regarding our example the total RUE can be derived by calculating all single values as shown above and afterwards taking the sum over the respective products which is not shown in detail due to the limited scope of this paper.

³Leaf nodes are not succeeded by arguments and RUE can only be determined for their parents.

4. Conclusions and Outlook

The purpose of this work is to present to our knowledge the first approach to calculate reflective user engagement in an Argumentative Dialogue System. Given a bipolar argumentation graph and fitting dialogue model, we propose a derivation which takes the depth, balance and number of inquiries into account.

In future work, we want to test the calculated RUE with simulated and real user data and explore its suitability for RL. Our aim is to cooperatively provide as much balanced information as possible, while adapting the system's strategy to the RUE.

Acknowledgments

This work has been funded by the DFG within the project "How to Win Arguments – Empowering Virtual Agents to Improve their Persuasiveness", Grant no. 376696351, as part of the Priority Program "Robust Argumentation Machines (RATIO)" (SPP-1999).

References

- [1] Annalena Aicher, Niklas Rach, Wolfgang Minker, and Stefan Ultes. Opinion building based on the argumentative dialogue system bea. *Increasing Naturalness and Flexibility in Spoken Dialogue Interaction: 10th IWSDS*, pages 307–318, 2021.
- [2] Ioannis Arapakis, Mounia Lalmas, and George Valkanas. Understanding within-content engagement through pattern analysis of mouse gestures. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, New York, NY, USA, 2014. Association for Computing Machinery.
- [3] Georg Buscher, Andreas Dengel, and Ludger van Elst. Query expansion using gaze-based feedback on the subdocument level. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, page 387–394, 2008.
- [4] Georges Dupret and Mounia Lalmas. Absence time and user engagement: Evaluating ranking functions. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, New York, NY, USA, 2013. Association for Computing Machinery.
- [5] Fiona Farr and Elaine Riordan. Students' engagement in reflective tasks: an investigation of interactive and non-interactive discourse corpora. *Classroom Discourse*, 3(2):129–146, 2012.
- [6] Joseph Grafsgaard, Joseph B Wiggins, Kristy Elizabeth Boyer, Eric N Wiebe, and James Lester. Automatically recognizing facial expression: Predicting engagement and frustration. In *Educational Data Mining 2013*, 2013.
- [7] Barbara A. Greene. Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist*, 50(1):14–30, 2015.
- [8] Qi Guo and Eugene Agichtein. Beyond dwell time: Estimating document relevance from cursor movements and other post-click searcher behavior. In *Proceedings of the 21st International Conference on World Wide Web, WWW '12*, page 569–578, New York, NY, USA, 2012. Association for Computing Machinery.

- [9] A. Haim and Oren Tsur. Open-mindedness and style coordination in argumentative discussions. In *EACL*, 2021.
- [10] Jeff Huang, Ryen White, and Georg Buscher. *User See, User Point: Gaze and Cursor Alignment in Web Search*, page 1341–1350. Association for Computing Machinery, New York, NY, USA, 2012.
- [11] Steve Jackson. *Cult of Analytics: Driving online marketing strategies using web analytics*. Routledge, 2009.
- [12] Siu Cheung Kong and Yanjie Song. An experience of personalized learning hub initiative embedding byod for reflective engagement in higher education. *Computers & Education*, 88:227–240, 2015.
- [13] M. Lalmas, H. O’Brien, and Elad Yom-Tov. Measuring user engagement. In *Measuring User Engagement*, 2014.
- [14] Janette Lehmann, Mounia Lalmas, Georges Dupret, and Ricardo Baeza-Yates. Online multitasking and user engagement. In *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, CIKM ’13*, page 519–528, New York, NY, USA, 2013.
- [15] Nona Lyons. Reflective engagement as professional development in the lives of university teachers. *Teachers and teaching*, 12(2):151–168, 2006.
- [16] Mohamad Noorman Masrek, Mohammad Hudzari Razali, Ishak Ramli, and Trias Andromeda. User engagement and satisfaction: The case of web digital library. *International Journal of Engineering and Technology (UAE)*, 7(4):19–24, 2018.
- [17] Maurizio Mauri, Pietro Cipresso, Anna Balgera, Marco Villamira, and Giuseppe Riva. Why is facebook so successful? psychophysiological measures describe a core flow state while using facebook. *Cyberpsychology, Behavior, and Social Networking*, 14(12):723–731, 2011.
- [18] Heather L. O’Brien and Elaine G. Toms. What is user engagement? a conceptual framework for defining user engagement with technology. *JASIST*, 59(6):938–955, 2008.
- [19] Jeeyun Oh, Saraswathi Bellur, and S. Shyam Sundar. Clicking, assessing, immersing, and sharing: An empirical model of user engagement with interactive media. *Communication Research*, 45(5):737–763, 2018.
- [20] Eli Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.
- [21] Olga Perski, Ann Blandford, Claire Garnett, David Crane, Robert West, and Susan Michie. A self-report measure of engagement with digital behavior change interventions (DBCI): development and psychometric evaluation of the “DBCI Engagement Scale”. *Translational Behavioral Medicine*, 10(1):267–277, 03 2019.
- [22] Ashok Kumar Ponnuswami, Kumaresh Pattabiraman, Qiang Wu, Ran Gilad-Bachrach, and Tapas Kanungo. On composition of a federated web search result page: Using online users to provide pairwise preference for heterogeneous verticals. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM ’11*, page 715–724, New York, NY, USA, 2011. Association for Computing Machinery.
- [23] Mario Alfonso Prado-Romero, Alberto Fernández Oliva, and Lucina García Hernández. Identifying twitter users influence and open mindedness using anomaly detection. In Yanio Hernández Heredia, Vladimir Milián Núñez, and José Ruiz Shulcloper, editors, *Progress in Artificial Intelligence and Pattern Recognition*, pages 166–173, Cham, 2018. Springer

International Publishing.

- [24] Gloria Jean Rodman. Facilitating the teaching-learning process through the reflective engagement of pre-service teachers. *Australian Journal of Teacher Education*, 35(2):20–34, 2010.
- [25] Chaklam Silpasuwanchai, Xiaojuan Ma, Hiroaki Shigemasu, and Xiangshi Ren. Developing a comprehensive engagement framework of gamification for reflective learning. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pages 459–472, 2016.
- [26] Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, and Suju Rajan. Beyond clicks: Dwell time for personalization. In *Proceedings of the 8th ACM Conference on Recommender Systems, RecSys '14*, page 113–120, New York, NY, USA, 2014. Association for Computing Machinery.