

# Explaining an Argumentation. Differences and Structural Analysis as a Foundation to Improve Case-Based Explanation

Jakob Michael Schoenborn,<sup>1</sup> Klaus-Dieter Althoff<sup>2</sup>

**Abstract:** The awareness for explanation-aware computing rose during the last two years rapidly. Because of the rising need in providing an explanation to support the decision made by an autonomous process, the amount of discussions and investigations on what is a “good” explanation increases as well. A differentiation between an explanation and an argumentation is missing. These two terms are often used synonymously, but are often aiming towards two different goals and thus, at a closer look, are structured differently. It has yet to be determined, if this structure can be used to identify “good” explanations and which structures the targeted conversational partner appreciates the most to increase the overall satisfaction.

**Keywords:** Explanation; Argumentation; Case-Based Explanation

## 1 Introduction

A recent case study provided by Binns et al [Bi18] has compared different styles of explanations regarding justice in algorithmic decision making: Input influence-based explanation, demographic-based explanation, case-based explanation and sensitivity-based explanation. Using these different styles of explanations, case-based explanation has been viewed as the least appropriate style with the most negative impact on justice perception [Bi18]. A case-based explanation supported the given decision by referring to “*thousands of similar cases from the past*” [Bi18] and illustrates the result using one exemplary case of the retrieved cluster of cases. This result seems surprising, given that case-based explanation is paired with case-based reasoning and thus is representing actual, experienced cases whereas e. g. demographic-based explanation might infer attributes to a person which are inappropriate. But by pointing out a single case, participants of the case study distanced themselves from the proposed case: “*This might be the case for x, but that does not mean it will happen to me.*” [Bi18]. In this case, the explanation was not specific to the user but rather a generic, broad statement. A possible solution to this problem is to discover a structure for an explanation that incorporates the users needs and state of knowledge. The user needs to be convinced, which ultimately results in changing one of the beliefs

<sup>1</sup> Intelligent Information Systems, University of Hildesheim German Research Center for Artificial Intelligence (DFKI) schoenb@uni-hildesheim.de

<sup>2</sup> Intelligent Information Systems, University of Hildesheim German Research Center for Artificial Intelligence (DFKI) klaus-dieter.althoff@dfki.de

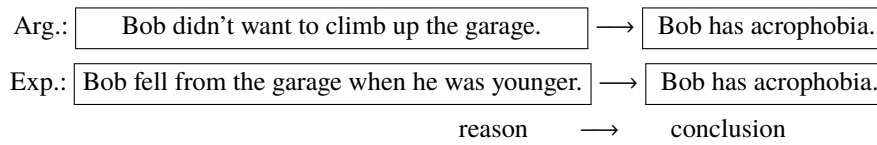


to a new or corrected belief [Mo02]. But to convince someone, an argumentation is the intuitive approach. In fact, argumentation and explanation are often used as exchangeable terms [LDA09, AWW09, CDL03, vLvdBtC18]. These terms have to be differentiated to distinguish an explanation from an argumentation, since studies have proved that an explanation is favored over an argumentation [Ku91, Ku01].

## 2 Argumentation and Explanation

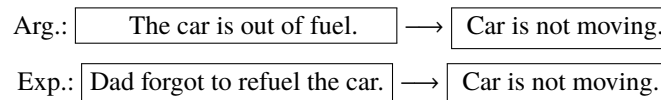
### 2.1 Differentiation and structure

One proposed way to distinguish an argumentation from an explanation is to look at which question has been addressed. A “how” question usually results in an argumentation while a “why” question results in an explanation:



While the upper conclusion answers the question “**How do you know**, whether Bob has acrophobia? Because Bob didn't want to climb up the garage.”, the lower conclusion answers “**Why** does Bob has acrophobia? Because Bob fell from the garage when he was younger” but the answers are not exchangeable: “**Why** does Bob has acrophobia? Because Bob didn't want to climb up the garage” does not seem to be an satisfactory explanation and might be an overeager assumption to infer acrophobia out of a denied activity. Furthermore, the explanation additionally answers the how-question here: “**How do you know**, whether Bob has acrophobia? Because Bob fell from the garage when he was younger”. In this special, artificial case, the explanation should always be preferred over an argumentation.

But this situation is not always the case. A counterexample would be:



Here, argumentation (the former) and explanation (the latter) are exchangeable since both questions can be answered using either argumentation or explanation. Depending on the users experiences, the user prefers a rather technical explanation (out of fuel) over a rather social explanation (Dad forgot to refuel the car). In this situation, it is merely possible for a knowledge engineer to design the correct explanation for the user, since there is no initial information whether the user wants to be convinced or is rather looking for helpful assistance.

Hence, the following definitions are used to distinguish between an argumentation and an explanation:

**Definition 1:** *An argument is a reason in which the fact functions as evidence in support of the conclusion. Its goal is to convince the conversational partner on the validity of the conclusion.*

**Definition 2:** *An explanation is a supportive, personalized information on top of a provided conclusion. In contrast to an argumentation, its goal is to help the conversational partner in understanding the reasoning behind the conclusion and its outcome.*

## 2.2 Explanations: Answers to why-questions

As motivated in the example above, an explanation can function as an argumentation, but not vice-versa. Thus, if the constructional overhead is reasonable, the knowledge engineer should always aim at providing an explanation rather than an argumentation. But why is it important to provide an explanation and when is it beneficial? It fosters trust into the system [GMW08, Sc94]. It might be arguable that an argumentation does achieve this goal as well, given its fact-based and fact-supportive nature, but a personalized explanation is more likely to be accepted [Mi19, Bi18]. As a user keeps using a certain system, the trust into that system increases by gaining routine through an increased amount of knowledge on how to use the system. If trust and explanation seems to be connected by the usage of a system, the question remains when it is actually useful to provide an explanation. This question has been answered by D. Leake who pointed out the importance of providing an explanation during an unexpected situation which “*must also identify the flaws in the understander’s prior beliefs that led it to generate flawed expectations.*” [Le95, p. 410]. These explanations are presented as explanation patterns (XP) as suggested by Schank [Sc94] which are basically deducing an explanation out of a set of given premisses. These patterns are limited due to their schema-based nature (and thus not being able to deal with novelty), but adding case-based reasoning provides an opportunity to point out these novel cases and provide an appropriate solution [Le95].

In contrast to an argumentation, an explanation can be issued on very different levels. Toulmin presented an legal argumentation structure, where an argumentation can be broken down to given data which qualifies a claim [To03]. This data is supported by warrants which are backed up, but these can be attacked by a rebuttal (see Fig. 1). While an argument is under attack, another argument has to be brought up to counter the attack. If the attack cannot be countered, the argument is considered to be invalid. If no further attack can be issued, the argument is valid. To follow these rules, an argument has to be issued in a measurable, fact-based way to leave no room of interpretation open. In an argumentation, there is no

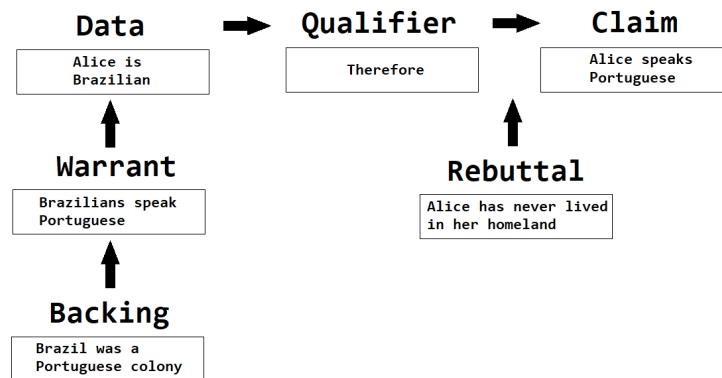


Fig. 1: Toulmin: Legal Argumentation Structure [To03] (example analogue to [Mo02])

acceptance for subjective statements since they aim to convince the conversational partner and this is done by objective arguments. This can be different for an explanation.

As defined above, an explanation aims to help the conversational partner by understanding the flaws in the prior beliefs. Thus, it is also possible to provide a subjective explanation. To clarify, an argumentation can be viewed as a subset of explanations: The informations used in an explanation can be categorized into multiple trust categories: On the highest level, the information is completely trusted (e. g. facts and statements issued by the conversational partner in question). This level might also be used by an argumentation. On lower levels, there are only assumptions, i. e. based on persons actions (Bob did not want to climb up the garage) or on a persons past (Bob fell off the garage when he was younger). Since these are only assumptions, these information might not be true during the past - and thus would be immediately attacked in the legal argumentation structure - but could be the last missing piece for the conversational partner to correct the flaw in the prior beliefs.

Lim et al. experimented on using why- and why-not explanations. During a case study, the participants were offered an explanation which basically tried to predict the correct answer whether a given person is exercising or not, based on parameters, e. g. body temperature and pacing [LDA09]. Nevertheless, the hypothesis to improve the user experience over having no explanation could be proven by a small margin. The same situation could also be observed during the studies of [CDL03, vLvdBtC18, Aa93, On97]. Even an increase by a small margin is valuable. Given the opportunity to foster user-centric explanations due to the growth in usage and techniques of NLP, it seems promising to revisit providing explanations. Another promising approach is the compositional adaptation of explanations in textual case-based reasoning [SÖM16] by extracting explanations from aviation incident reports and store them as text reasoning graph. Nodes in the text reasoning graph are extracted phrases and sentences while edges correspond to causal and entailment relations:

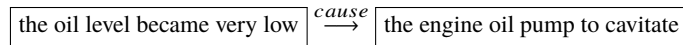


Fig. 2: Part of the text reasoning graph provided by Sizov et al., translated into the given definition of an explanation [SÖM16]

Even though the user experience has been increased by providing these explanations, the authors pointed a few erroneous conclusions out due to more distant cases being reused [SÖM16]. Since this might seem to be a maintenance issue, it could also be possible to counteract this problem by providing a solid structure for an explanation.

### 3 Future work

As pointed out before, the next step is to identify possible structures to increase the benefit of using case-based reasoning and to store/reuse explanations in a more efficient manner. Another possible foundation to build up on a structure for explanations might be the usage of the SIAM methodology provided by T. Roth-Berghofer [Ro03]. This allowed to overcome the limitation of Toulmins legal argumentation structure and iterate over the mentioned explanations with a lower level of trust. Furthermore it is crucial to identify the users needs by classifying the current state of knowledge to provide an user-specific explanation, and when an explanation is needed beyond the occurrence of unexpected events. It has also to be reviewed, in which cases an argumentation is sufficient to correct help the user and when actually an explanation with its necessary construction overhead has to be build up. This is important, because justification (which is one of the most important attributes of an explanation as pointed out by multiple authors [SCA05, Ei18, Li11]) can be the most suitable factor to increase the users satisfaction.

### References

- [Aa93] Aamodt, Agnar: Explanation-Driven Retrieval, Reuse and Learning of Cases. In: University of Kaiserslautern (Germany). pp. 279–284, 1993.
- [AWW09] Awad, Ahmed; Weidlich, Matthias; Weske, Mathias: Specification, Verification and Explanation of Violation for Data Aware Compliance Rules. In (Baresi, Luciano; Chi, Chi-Hung; Suzuki, Jun, eds): Service-Oriented Computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 500–515, 2009.
- [Bi18] Binns, Reuben; Van Kleek, Max; Veale, Michael; Lyngs, Ulrik; Zhao, Jun; Shadbolt, Nigel: 'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, ACM, New York, NY, USA, pp. 377:1–377:14, 2018.
- [CDL03] Cunningham, Pádraig; Doyle, Dónal; Loughrey, John: An Evaluation of the Usefulness of Case-Based Explanation. 2003.

- [Ei18] Eisenstadt, Viktor; Espinoza-Stapelfeld, Christian; Mikyas, Ada; Althoff, Klaus-Dieter: Explainable Distributed Case-based Support Systems: Patterns for Enhancement and Validation of Design Recommendations. 07 2018.
- [GMW08] Glass, Alyssa; McGuinness, Deborah L.; Wolverson, Michael: Toward Establishing Trust in Adaptive Agents. In: Proceedings of the 13th International Conference on Intelligent User Interfaces. IUI '08, ACM, New York, NY, USA, pp. 227–236, 2008.
- [Ku91] Kuhn, Deanna: *The Skills of Argument*. Cambridge University Press, 1991.
- [Ku01] Kuhn, Deanna: How do People Know? *Psychological Science*, 12(1):1–8, 2001. PMID: 11294222.
- [LDA09] Lim, Brian Y.; Dey, Anind K.; Avrahami, Daniel: Why and Why Not Explanations Improve the Intelligibility of Context-aware Intelligent Systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09, ACM, New York, NY, USA, pp. 2119–2128, 2009.
- [Le95] Leake, David B.: Abduction, experience, and goals: a model of everyday abductive explanation. *J. Exp. Theor. Artif. Intell.*, 7(4):407–428, 1995.
- [Li11] Lillehaug, Marvin B.: *Explanation-aware case-based reasoning*. 2011.
- [Mi19] Miller, Tim: Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1 – 38, 2019.
- [Mo02] Moulin, Bernard; Irandoust, Hengameh; Belanger, Micheline; Desbordes, Gaele: Explanation and Argumentation Capabilities: Towards the Creation of More Persuasive Agents. *Artif. Intell. Rev.*, 17:169–222, 05 2002.
- [On97] Ong, Lean Suan; Shepherd, Barry; Tong, Loong Cheong; Seow-Choen, Francis; Ho, Yik Hong; Tang, Choong Leong; Ho, Yin Seong; Tan, Kelvin: The colorectal cancer recurrence support (CARES) System. *Artificial Intelligence in Medicine*, 11(3):175 – 188, 1997.
- [Ro03] Roth-Berghofer, Thomas: Knowledge maintenance of case-based reasoning systems - the SIAM methodology, volume 262 of DISKI. Infix Akademische Verlagsgesellschaft, 2003.
- [Sc94] Schank, Roger C.: *Inside Case-Based Explanation*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1st edition, 1994.
- [SCA05] Sørmo, Frode; Cassens, Jörg; Aamodt, Agnar: Explanation in Case-Based Reasoning—Perspectives and Goals. *Artificial Intelligence Review*, 24(2):109–143, Oct 2005.
- [SÖM16] Sizov, Gleb; Öztürk, Pinar; Marsi, Erwin: Compositional Adaptation of Explanations in Textual Case-Based Reasoning. In: ICCBR. volume 9969 of Lecture Notes in Computer Science. Springer, pp. 387–401, 2016.
- [To03] Toulmin, Stephen E.: *The Uses of Argument*. Cambridge University Press, 2 edition, 2003.
- [vLvdBtC18] van Loon, Maria; van den Broek, Sjoukje; ten Cate, Olle: A Model Study Guide for Case-Based Clinical Reasoning. In (ten Cate, Olle; Custers, Eugène J.F.M.; Durning, Steven J., eds): *Principles and Practice of Case-based Clinical Reasoning Education : A Method for Preclinical Students*. Springer International Publishing, Cham, pp. 121–132, 2018.