Evidence based Semantics for Reasoning beyond Your Data

Yukio Ohsawa [1], Kaira Sekiguchi[1], Tomohide Maekawa [2], Hiroki Yamaguchi[2] , and Sae Kondo[3]

[1] *The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
[2] *Trust Architecture Inc., 509 Louis Marble Nogizaka, 9-6-30 Akasaka, Minato-ku, 107-0052,Tokyo*
[3] *Mie University, Kurimamachiya-cho Tsu city, Mie 514-8507, Japan*

Abstract

We present a new problem for the society of AI, that is to collect evidence useful for explaining the meaning of an observed event. Evidence here is a set of pieces of useful and new information, which may come from the open real world, for understanding the target observation. The information may not be necessarily accepted widely enough to be learned by machine learning but is novel knowledge or claim within personal or local messages, which supports a query positively or negatively. We address this presentation to the proposal of Evidence-based Semantics, which means to obtain a novel explanation of the meaning of an event, situation, action, utterance, message, etc., that is critically required in various real-world application domains.

Keywords

Evidence collection, Semantics, Abduction, Visualization, Logic Tree

1. Introduction

Learning reusable patterns, extracting useful parts of data, and searching for data hitting user's interest are already daily tasks for a machine. By combination, it is easy to collect datasets and use them for knowledge discoveries if the user can express his/her own interest. The recent Chat GPT, where the user enters the query in natural language, returns an answer which may appear to come from the machine's understanding of the query and relevant information in collected various datasets.

However, it is still an open problem to respond to quite a simple and widely perceived requirement, that is to collect useful information from the open space of information for explaining the meaning of a query i.e., an event observed in the real life. The required information may not be collected from the open data or public data market processed by machine learning, but may be novel knowledge or claims within personal or local messages. We call this kind of information *evidence* and address this short paper to the proposal of Evidence-based Semantics, which means to explain the meaning of a target observation (an event, situation, action, message, etc) by collecting evidence. In contrast to the literature [1, 2], we seek evidence for creating hypotheses for explaining the meaning of an event rather than for labeling (T/F) hypotheses which are given or generated from given hypotheses.

Evidence-based Semantics

EBS refers to the problem to obtain E and h', given Σ and h as in Clauses (1) through (3), where G and Σ respectively refer to the target observation and prepared knowledge. Σ can be probabilistic, but it is consistent with new knowledge or observation.

$$\Sigma \cup h \quad \vdash G \tag{1}$$

$$\Sigma \cup h' \quad \vdash G \cup E \tag{2}$$

$$h \neq h' \tag{3}$$

$$\Sigma \cup h' \quad \nvdash \square \tag{4}$$

h, and h', are respectively hypotheses to entail G, and G together with new observation E. E is the evidence. a collection of additional observations (e_i's in Fig.1) supporting h'. G can be entailed by the previous hypothesis h, but the new hypothesis h' is preferred due to its ability to entail both G and E.

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27–29, 2023, USA
EMAIL:info@sys.t.u-tokyo.ac.jp (A.1); tmaekawa@trst-arc.com (A.2), skondo@arch.mie-u.ac.jp (A. 3)
ORCID: 0000-0003-2943-2547 (A.1); 0000-0001-6473-3391(A.2); 0000-0003-2162-0287 (A.3)

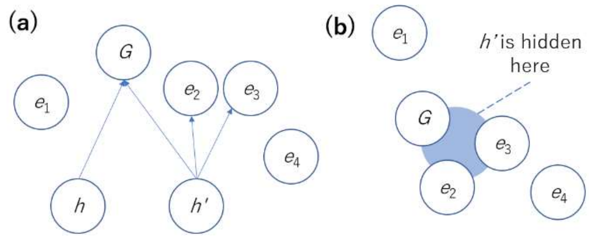

Fig. 1. Two approaches to EBS. In (a), e_i's in solid circles, entailed by the hypothesis h' (h_3) entailing G are elements of E. In (b), events, not hypotheses, are visualized on data.

First level heading

Two approaches to EBS are shown in Figure 1, where (a) represents the entailment structure above, where e and h are observed events other than G and hypotheses respectively. The arrows show parts Σ used for explaining (entailing) G and e. By use of data visualization in (b), h's are invisible (not observed) but a hypothesis of higher confidence is reflected here as one at a shorter distance from observations derived from it. Here we regard these two figures as the bases of approaches toward EBS.

Approach A: Logic Trees

Fig. 2 shows the process using dFrome [3] to extend the goal (a) given on a logical flow about an imaginary accusation of a conflict in a workplace, to which relevant pieces of knowledge are extracted from the logic-tree base and added as in (b). Causal events are thus searched and added to finally obtain (c), which corresponds to Fig.1 (a).

Approach B: Data Visualization

Using the multi-layer KeyGraph [4] as in Fig.3, the foreground graph in the center represents an abstract of an imaginary accusation similar to the case in Fig.2, the background messages in the workplace. The messages (e.g. "sec10") close to the foreground graph came to be evidence to explain the meaning of supervisor's behaviors, corresponding to Fig.1 (b).

Conclusions

EBS can be regarded as a restoration of the situational semantics [5] in the recent context of data exchange and utilization. In comparison with the classical logical abduction as in [6], in EBS, the initial hypothesis is challenged by more consistent hypothesis via observations in the open world. In the next step, we are developing a method to explore and link data in the open world to tools for EBS.

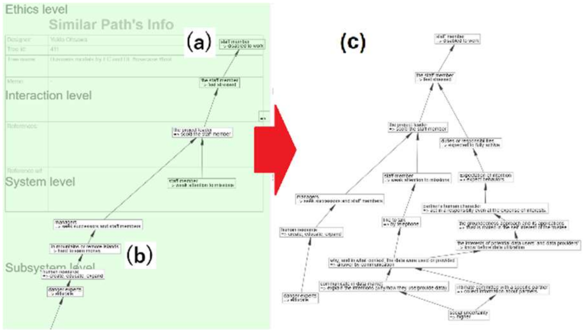

Fig. 2. A sequence of goal-directed logical flows, importing other flows by social coupling via similarities between words

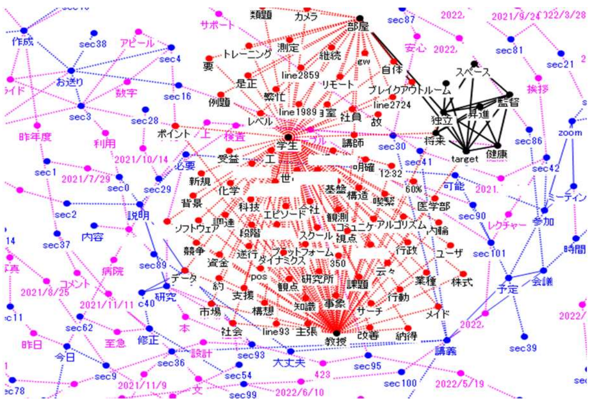

Fig. 3. Visualizing text relevant to an imaginary accusation

Acknowledgement

This study is supported by JST COI-NEXT JPMJPF2013, MEXT Q-Leap JPMXS0118067246, JSPS Kakenhi 20K20482. We thank Lawyer R. Uchino for advices about evidence collection in disputes and statements in this paper.

[1] Gupta, V., et al, Right for the Right Reason: Evidence Extraction for Trustworthy Tabular Reasoning, *Proc. Ann. Meet. Assoc. for Comp. Ling.* 1., 3268 – 3283 (2022)

[2] Mokhtar, S.H., et al, A Review of Evidence Extraction Techniques in Big Data Environment, in *Proc. ICSCEE* (2018)

[3] Sekiguchi, K., Hori, K. Organic and Dynamic Tool for Use with Knowledge Base of AI Ethics for Promoting Engineers' Practice of Ethical AI Design, *AI & Soc.* 35, 51-71, 2020.

[4] Ohsawa, Y., Benson, N.E., Yachida, M., KeyGraph: automatic indexing by co-occurrence graph based on building construction metaphor, *Proc. IEEE ADL*'98, CA, USA, 1998, pp. 12-18

[5] Barwise, J., Perry, J. *Situations and Attitudes,* MIT Press (1983).

[6] Poole, D., Probabilistic Horn abduction and Bayesian networks, *Artif. Intell.* 64 (1), 81-129 (1993)