The light and dark side cohabiting the same graph: knowledge patterns behind the 'half-true' fact-checks

Kateřina Haniková[1], David Chudán[1], Vojtěch Svátek[1], Jana Syrovátková[1] and Marek Dudáš[1]

[1]*Faculty of Informatics and Statistics, Prague University of Economics and Business, nám. W. Churchilla 1938/4, Prague, Czech Republic*

Abstract

Fact-checking is an activity aiming to establish and explain the veracity of salient claims appearing in the public space. Of particular importance are the claims that are deemed half-true, as they might most easily mislead even attentive and well-informed readers. In this early-phase research, we attempted to interpret the argumentation of fact-checkers for such claims as descriptions of situations consisting of objects and their relationships and valuations, parts of which are indicated as false or misleading. Based on a graph-based and/or semi-informal textual representation of a sample of fact-checks, we aimed to distill recurring knowledge patterns having simple ontology structures as their components. This could, in turn, support NLP tasks such as argumentation extraction or fact-check summarization.

Keywords

fact-checking, argumentation structure, knowledge pattern, falsehood annotation

1. Introduction

Fact-checking is nowadays a topical activity, aiming at determining the veracity of statements posted by influential people both in mainstream media and in social media. The fact-checkers analyze the background of individual statements (called claims in this context), and write up reports or articles in which they not only provide the verdict (whether the claim is true, false, half-true, unsubstantiated, or the like), but also elaborate argumentation. This argumentation involves concrete entities (people, organizations, dates, money amounts, etc.), and also their relationships, some of which may be labeled as true and some as false, but also misleading or unsubstantiated. We hypothesize that these argumentation structures can be mapped on (a specific kind of) knowledge patterns, and are interested in exploring the frequencies of such patterns in different sets of fact-checks.

In the rest of the paper, we first present five example argumentation graphs based on real fact-checks, then proceed to a (small-scale, single-source, and necessarily rough) empirical analysis of the subpatterns expressing the 'dark side' of the claims. Finally, we provide a few links to mildly related prior research, and wrap up the paper including future research.

14th Workshop on Ontology Design and Patterns (WOP) 2023, Athens, Greece

katerina.hanikova@vse.cz (K. Haniková); david.chudan@vse.cz (D. Chudán); svatek@vse.cz (V. Svátek); jana.syrovatkova@vse.cz (J. Syrovátková); marek.dudas@vse.cz (M. Dudáš)

CEUR Workshop Proceedings (CEUR-WS.org)

2. Fact-check graph examples in PURO

As a starting point, we manually analyzed a collection of fact-checks from the *politifact.org* site. We focused on 'half-true' fact-checks,[1] assuming that their argumentation graphs can be richer than those of mere true or false fact-checks, and also considering their topicality, as half-true claims can possibly mislead the readers the most. The analysis proceeded in two steps; both were carried out by a knowledge engineer with discourse analysis erudition.

The first step resulted in a *semi-informal* textual structure, already pointing out the entities (objects), their relationships and attributes that together form the false/misleading aspects of the claim, and also the true statements that are used either in the claim itself (possibly even consciously, to make the whole statement more trustworthy) or in the fact-check (to prove the fallacy of the claim by contrasting the invalid statements). In most cases, all necessary information was present in the fact-check summary part.

In the second step, we leveraged on both our semi-informal summary and the original politifact summaries (resorting to the full text where needed), and created a *graph representation* of the situation. For this purpose we used *PURO Modeler*, a web-based graphing tool conforming to the PURO language [1]; the latest version of the tool is available as a demo.[2] The adequacy of PURO Modeler for this task follows from the fact that:

- It has distinct built-in constructs for particulars (instances) vs. universals (types), for relationships vs. objects, and also for quantitative-valued attributes.It thus allows to easily draft instance knowledge graph structures together with their types, and also quantitative values, which are crucial in many fact checks.
- Since it is more relaxed than graph editors conforming to the OWL-DL metamodel (it allows for higher-order types and for higher-arity and higher-order relationships), it allows to defer some encoding decisions (such as metamodeling or classes by instances or n-ary relation reification) that might have to be made in an OWL-DL graph editor.

The PURO primitives are represented, in the graph, by rectangles (objects), ovals (types), relationships (diamonds) and quantitative attributes (hexagons). As regards the explicit labeling of the false/misleading claim aspects in the graph, since PURO Modeler does not provide customizable graphical features (e. g., changing the color of nodes), we decided to reuse the *relationship* construct (green diamond) for annotating such graph elements, while distinguishing such annotations from regular relationships by employing *capitalized* labels. The graphs do not capture the whole argumentation structure, but rather just the part labeled as false and its close neighborhood. We also want to make clear that the 'falsehood' annotations do not express any formal semantics over the graph entities (they are in fact the least rigorous part of the graphs).

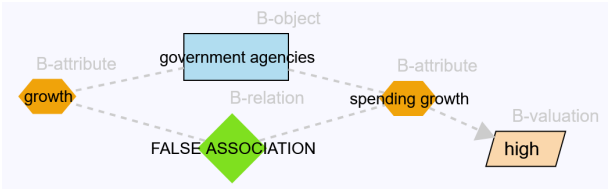
In Fig. 1 we show five example graphs. They refer to the five politifact 'half true' fact-checks published between June 13 and July 13, 2023; the fact-checks and their summaries can be retrieved from the politifact website. We will now comment on the graphs one by one.

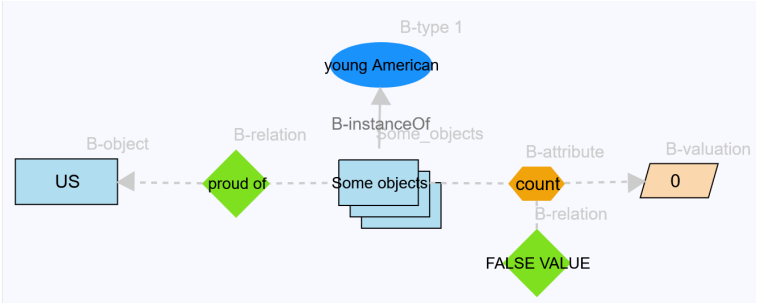
The first[3] example addresses a claim that used the figure on government agencies' total spending growth to demonstrate how the agencies *themselves* grow in size. The spending

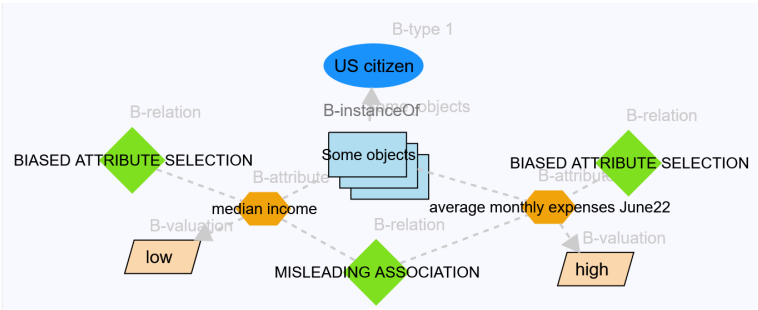
[1] https://www.politifact.com/factchecks/list/?ruling=half-true

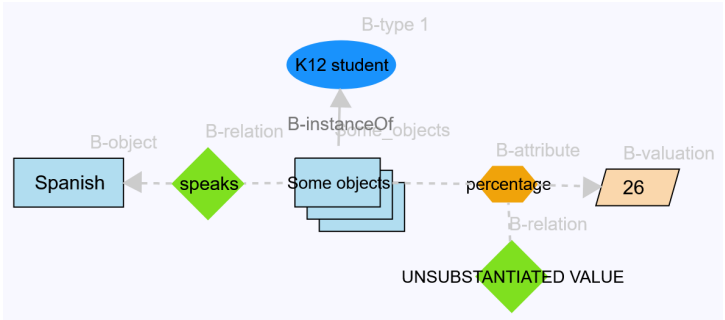
[2] https://protegeserver.cz/purom5/

[3] https://www.politifact.com/factchecks/2023/jul/13/ron-desantis/ron-desantis-statement-federal-government-growth/









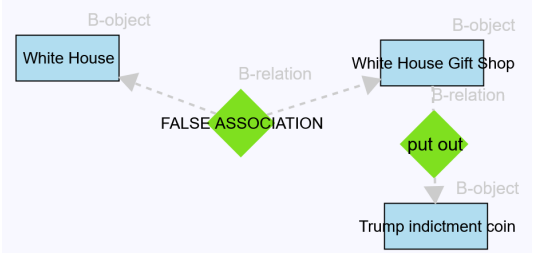


Figure 1: PURO models of five politifact fact-checks

however larger consists of payments used for citizens rather than for the agencies themselves. Therefore, we link the two attributes with a 'FALSE ASSOCIATION' annotation.

The second[4] example addresses a claim that is false if we consider it literally: that no young Americans are proud of being Americans (while, by the polls, some are still significantly proud, and a majority of them are proud to some degree). We use here a specific primitive of PURO Modeler called 'Some objects' and denoting an unspecified number of objects for which some common feature/s (types and/or relations) hold.[5]

The third[6] example is more complex. In fact, it does not contain any clear non-truth, but the two attributes (a non-standard kind of median income measure, and the average monthly expenses for an untypical time period) presented together were specifically chosen in order to suggest a misleadingly strong contrast between the income and cost sides of US citizens' budget. The problematic aspects of the claim are thus dispersed between the mere choice of the attributes and their association through their synchronous presence in the claim.

The fourth[7] example is similar to the second one in terms of containing a 'Some objects' node in connection with a numerical value. The difference is that now the claim refers to a percentage rather than to the count (here, of a certain category of students), and, that the fact-checker did not explicitly refute the veracity of the percentage but pointed out that this particular percentage had been officially published for a slightly different measure (of ethnic Hispanics rather than of speakers of Spanish). It is unlikely but not completely impossible that the values of the two measures would coincide; therefore we use annotate the value as 'unsubstantiated' rather than as 'false'.

The fifth[8] example features a false association similar to the first one. However, the improperly coupled graph elements are objects rather than attributes this time: the claim tacitly suggests that a business entity (that has put out a controversial coin) called White House Gift Shop is officially affiliated to the government, while this is not the case (any longer).

In this initial phase of the research, the graph design was guided by parsimony principles. In particular, the following structures were intentionally omitted:

- True facts that, by the guess of the fact-checker, the authors of the claim may have used as basis/inspiration for their false statements (but they do not refer to those facts explicitly). This holds, e.g., for the percentage of Hispanic students in the fourth example.
- True facts the fact-checker used to disprove the false facts: e.g., the statistics of young people proud to be Americans, in the second example.
- Attributes that would be free of bias, in contrast to the biased ones, as in the third example.
- Situation in the past under which the claim was actually true, as in the fifth example (since the White House Gift Shop used to be an official site of White House gifts).

These additional structures are however documented in the semi-informal texts resulting from the first step of analysis. For the given five examples, they are available in an auxiliary page.[9]

[4] https://www.politifact.com/factchecks/2023/jun/28/vivek-ramaswamy/vivek-ramaswamy-is-partially-right-about-patriotis/
[5] The primitive was included upon a study of the so-called MISO ("multiple indirectly specified objects") family of ontology patterns [2].
[6] https://www.politifact.com/factchecks/2023/jun/26/tim-scott/tim-scotts-misleading-statements-on-median-income/
[7] https://www.politifact.com/factchecks/2023/jun/21/joe-biden/joe-biden-overstates-the-number-of-k-12-students-w/
[8] https://www.politifact.com/factchecks/2023/jun/13/bill-cassidy/gift-shop-selling-coin-for-trumps-indictment-is-no/
[9] https://bit.ly/WOP23-politifact

3. Small empirical study on falsehood patterns

To complement this initial knowledge pattern 'peep' with a study on a larger sample from a different source, we randomly selected 50 fact-checks from the largest[10] Czech fact-checking site, *demagog.cz*. Since demagog.cz does not have a category called 'half-true', we sampled on its closest counterpart, 'misleading'. Similarly to the politifact case, argumentation structures were first distilled as semi-informal textual statements – now by three junior knowledge engineers (each for a pool of the sample), briefly trained for the task through a handful of examples. Then an experienced knowledge engineer leveraged on the semi-informal statements as well as the original fact-check and claim so as to conclude on a *falsehood pattern* (which would become a falsehood annotation in graph-based modeling of the argumentation). This yielded 52 pattern instances, since in two cases there were two different falsehood annotations present.

The most common pattern (18 times) was FALSE VALUE. The second most common (15 times) was a new pattern, FALSE RELATIONSHIP, covering situations such as a party falsely claimed to have proposed a certain law or two events falsely claimed to have happened at the same time. Another new pattern was MISSING RELATIONSHIP (4 times), e.g., in a case where an unmentioned relationship compensated the adverse effect of a relationship spotlighted in the claim. There were also UNSUBSTANTIATED RELATIONSHIP (3 times) and UNSUBSTANTIATED VALUE (4 times), which differed from their 'false' counterparts in the sense that the relationship/value could not be disproved by the fact-checker outright, but no evidence was available for them. In contrast to the politifact micro-sample, no misleading 'side-by-side' associations were found; it seems that demagog.org is geared towards claims that can be disproved at the *factual* level rather than towards pointing at unfair *propaganda techniques* in the claims. Mildly in the 'side-by-side' spirit is perhaps the pattern FALSE SIMILARITY (3 times), dealing with unsubstantiated claims that two objects (e.g., organizations or events) are very similar while they are not. The remaining five identified patterns only occurred once in the whole sample.

4. Related research

Formal modeling of argumentation is an area of extensive research. Of the ontology-based approaches, we should mention the Argument Model Ontology (AMO) [3], based on Toulmin's theory. Another attempt to formalize argumentation structures and to interchange data between argumentation tools is the Argument Interchange Format (AIF) [4].

The long history of argument diagramming is overviewed in a journal article [5]. One such diagramming tool is Araucaria [6],[11] which is based on the Argumentation Markup Language (AML). The main difference between our PURO diagrams and those created in Araucaria is that the argument diagram created using Araucaria is focused on the whole sentences, whereas diagrams designed in PURO depict entities and relationships between them. Overall, existing argumentation ontologies and modeling tools could be used for drawing argumentation maps based on fact-checks, however they fail to support entities and relationships in these argumentation structures, which is the focus of our study.

[10] And also the only accredited by the International Fact-Checking Network, see https://www.poynter.org/ifcn/.
[11] http://araucaria.arg.tech/doku.php?id=start

5. Conclusions and future work

We demonstrated how the structure of fact-checks can be to some degree captured through ontological diagrams of entities, relationships and attributes, and that often similar patterns recur in different fact-checks.

Future work will investigate whether the diagrams can speed up the understanding of the fact-check, and whether such speed-up depends on the particular patterns employed. We will also explore if and how the PURO diagrams of fact checks can be (using our existing transformation tool, OBOWLMorph [1], or another one) uniformly transformed to OWL/RDF, so that they could be integrated with other knowledge graphs related to fact-checking and argumentation in general. We will also try to align the graph structures of fact-checks both with ontology patterns from repositories such as *ontologydesignpatterns.org* and with established argumentation theories, and, consequently, set up a fact-check pattern catalog. Finally, in longer term, we would like to explore the use of machine learning techniques for identification of cataloged patterns in the texts of fact-checks, as well as for fact-check summarization, since a fact-check summary should 'spotlight' on the proof of falsehood, possibly in the spirit of our falsehood annotations in graphs.

Acknowledgments

The research is supported by CHIST-ERA within the CIMPLE project (CHIST-ERA-19-XAI-003).

References

[1] M. Dudáš, V. Svátek, M. Vacura, O. Zamazal, Starting ontology development by visually modeling an example situation – A user study, in: VOILA@ISWC 2016, Kobe, Japan, October 17, 2016, volume 1704, CEUR-WS.org, 2016. URL: https://ceur-ws.org/Vol-1704/paper10.pdf.

[2] V. Svátek, J. Kluka, M. Vacura, M. Homola, M. Dudás, Patterns for referring to multiple indirectly specified objects (MISO): analysis and guidelines, in: Advances in Pattern-Based Ontology Engineering, volume 51 of *Studies on the Semantic Web*, IOS Press, 2021, pp. 1–24. URL: https://doi.org/10.3233/SSW210004. doi:10.3233/SSW210004.

[3] Fabio Vitali, Silvio Peroni, The Argument Model Ontology, 2011. URL: https://sparontologies.github.io/amo/current/amo.html.

[4] C. Chesñevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk, S. Willmott, Towards an argument interchange format, The Knowledge Engineering Review 21 (2006) 293–316. URL: https://doi.org/10.1017/S0269888906001044. doi:10.1017/S0269888906001044.

[5] C. Reed, D. Walton, F. Macagno, Argument diagramming in logic, law and artificial intelligence, The Knowledge Engineering Review 22 (2007) 87–109. doi:10.1017/S0269888907001051.

[6] C. Reed, G. Rowe, Araucaria: Software for argument analysis, diagramming and representation, International Journal on Artificial Intelligence Tools 13 (2004) 961–979. URL: https://doi.org/10.1142/S0218213004001922. doi:10.1142/S0218213004001922.