Leveraging Educational Technology to Improve the Quality of Civil Discourse

Nicholas Diana(✉)

Carnegie Mellon University, Pittsburgh, PA 15213, USA
ndiana@cmu.edu

Abstract. The ability to critically assess the information we consume is vital to productive civil discourse. However, recent research indicates that Americans are generally not adept at, for instance, identifying if a news story is real or fake. We propose a three-part research agenda aimed at providing accessible, evidence-based technological support for critical thinking in civic life. In the first stage, we built an online tutoring system for teaching logical fallacy identification. In stage two, we will leverage this system to train crowd workers to identify potentially fallacious arguments. Finally, in stage three, we will utilize these labeled examples to train a computational model of logical fallacies. We discuss how our current research into instructional factors and Belief Bias has impacted the course of this agenda, and how these three stages help to realize our ultimate goal of fostering critical thinking in civil discourse.

Keywords: Cognitive tutors · Informal logical fallacies
Informal reasoning · Cognitive task analysis
Difficulty factors assessment · Ill-defined domains

1 Introduction

The recent rise of *fake news* on popular social media platforms has underscored the importance of critical thinking skills in informed civil discourse. Some estimates claim that Americans fall for fake news headlines approximately 75% of the time [5], and as reliance on social media as a primary new source increases, this problem may only get worse. One potential way to combat fake news is to embed support for key skills such as evidence evaluation and argument analysis into the media we consume. We propose that educational technology provides an opportunity for accessible, evidence-based instruction and support for these essential critical thinking skills, and describe our current and proposed research in this area.

2 Current and Proposed Work

Informal reasoning shares many of the challenges associated with other ill-defined domains (ambiguity, limited formal theories, etc.) [4]. To make the ill-defined

C. Penstein Rosé et al. (Eds.): AIED 2018, LNAI 10948, pp. 517–520, 2018.
https://doi.org/10.1007/978-3-319-93846-2_97

domain of informal reasoning more tractable, we focus on teaching informal fallacies. Informal fallacies are simply patterns of bad argumentation, where the premises of an argument fail to support the conclusion. More importantly, informal fallacies provide a semi-structured way to research how people reason about everyday arguments and the biases that may impact their logical reasoning. Our research agenda progresses towards this goal in three incremental stages.

2.1 Online Logical Fallacy Tutor

The first stage in exploring this space is to better understand how people learn to identify logical fallacies, and if educational technology can be an effective tool in teaching that skill. We have already taken some preliminary steps towards answering these questions by building and testing an online tutoring system designed to teach logical fallacy identification. The results of this initial foray are promising. First, our data suggest that a student's political beliefs may inhibit their ability to identify faulty logic in arguments that align with their own beliefs, a cognitive bias known as Belief Bias [3]. For example, liberal users have more difficulty identifying faulty logic in arguments with a liberal conclusion. This interaction between personal beliefs and our ability to identify faulty logic suggests that new solutions in this space may need to account for the user's beliefs in the instruction or intervention.

Our data also show clear evidence of learning. Specifically, we were interested in whether instruction that promoted deductive reasoning (e.g., definitions and explanations) was more effective than instruction that promoted inductive reasoning (e.g., seeing many examples of the fallacy). To test this, we conducted an Instructional Factors Analysis [2], and found that while users were able to learn a logical fallacy through inductive instruction (examples) alone, the deductive instruction (definitions and explanations) was more than twice as effective.

2.2 Crowdsourcing Logical Fallacy Detection

In addition to training citizens to identify logical fallacies themselves, we are also interested in ways that technology can support fallacy identification. It is likely the case that logical fallacies are easier to spot in the context of an online tutoring system simply because users are looking for them. In real life, we rarely expend so much energy thinking critically about each argument we read or hear (unless we fervently disagree with it), and real-world fallacies are often subtler than the illustrative examples given in textbooks [1]. Coupled with the pace of discourse in, for example, a presidential debate, logical fallacies are easy to miss, if youre not paying careful attention.

In this stage, we propose a system that helps users pay closer attention to specific pieces of discourse that might warrant some critical evaluation. We will leverage the online tutoring system we built previously to train or evaluate crowd workers. On the backend, our proposed system will break discourse content into digestible parts and then distribute the parts to crowd workers (trained using our tutoring system) who will flag any content that may contain a fallacy. On

the frontend, a user reading an article with our service enabled may notice a section that has been highlighted, indicating the possibility of faulty logic. It is important here to note that our crowd workers are not simply labeling arguments as fallacious, but rather as potentially fallacious. There are two reasons for this choice. First, the presence of a logical fallacy does not necessarily invalidate an argument. If I were to say, The earth is round and youre crazy if you dont think so. Id be committing an Ad Hominem fallacy, but the earth is still round. In many cases, arguments containing logical fallacies are not invalid, just weak arguments. Second, and more importantly, by not passing a verdict we force the user to exercise their own critical thinking skills. In this way, our system (1) provides more opportunities for thinking critically about real-world discourse, and (2) scaffolds those opportunities with some degree of judgement about the discourse in question.

We can validate the performance of crowd workers by comparing their classifications against the classifications of other crowd workers, or to a smaller set of expert classified examples. Evaluating the impact of the system on a users critical thinking skills is more difficult. One possibility is asking users to report if they agree that a fallacy exists in the highlighted text. Once we have validated the high likelihood of a fallacys existence in that piece of text (using the methods mentioned previously), we could use user agreement as a rough proxy for critical thinking ability. Furthermore, while we would expect agreement to be lower for arguments that align with the users political orientation (due to the aforementioned Belief Bias), we might expect that agreement would increase on these types of arguments over time, as critical thinking skills improve.

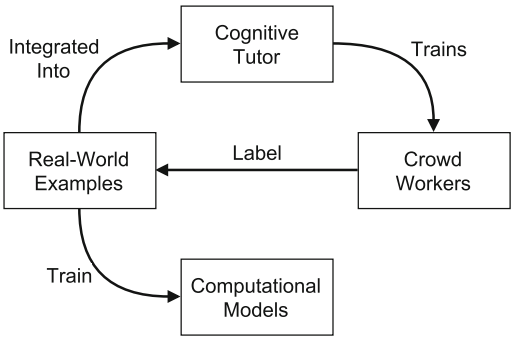

Fig. 1. Feedback loop relationship between the cognitive tutor and crowd workers. The real-world examples labeled by crowd workers can be used to both improve the cognitive tutor and train computational models. In addition to improving our tutoring system, we can also use those labeled examples to train a computational model of logical fallacies. We argue that machines may be able to recognize the possibility of a logical fallacy (using linguistic cues or similarity to prior labeled data). Once a potential fallacy is recognized, the machine will then prompt the human to assess the validity. This hybrid solution capitalizes on the computational strength of machines and the as-yet unparalleled power of human sense-making to improve fallacy identification.

2.3 Towards a Computational Model of Logical Fallacy Identification

The final goal of this work is to develop a computational model of logical fallacies. Achieving this goal requires overcoming several large challenges. First, to train a model to detect such a nuanced use of language will most likely require a large number of labeled examples. Furthermore, these examples will most likely have to be varied and authentic (perhaps unlike many of the purposefully illustrative examples used in textbooks). To solve this shortage of labeled examples, we propose using our cognitive tutor to train crowd workers to identify fallacies in real-world media sources (as described in the *Crowdsourcing* stage). High quality labels can then be automatically integrated into the tutor training system, increasing the number of potential examples crowd workers can use to achieve mastery. This increase in the number of examples may be especially important for fallacies that are most efficiently taught using inductive instruction. Figure 1 shows the feedback loop relationship between crowd workers and the cognitive tutor.

If we meet these challenges and are able to detect the markers of logical fallacies in real-world text, there are potential applications in media (both traditional and social), politics, and education. One could imagine a plugin for your favorite word processor that underlines a potential Appeal to Ignorance just as it would a misspelled word. Similarly, one could imagine how broadcasts of presidential debates in the future might be accompanied by a subtle notification anytime a candidate might be committing a Moral Equivalence fallacy.

Taken together, these projects represent an accessible, theory-driven way to impact the quality of civil discourse using educational technology. Just as technology has played an unfortunate role in amplifying fake news, so too will it play a role in combating those unreliable sources. We believe the most effective defense against this surge of misinformation are systems that both train citizens to be better critical thinkers and support that critical thinking in their everyday lives.

References

1. Boudry, M., Paglieri, F., Pigliucci, M.: The fake, the flimsy, and the fallacious: demarcating arguments in real life. Argumentation **29**(4), 431–456 (2015)
2. Chi, M., Koedinger, K.R., Gordon, G.J., Jordon, P., VanLahn, K.: Instructional factors analysis: a cognitive model for multiple instructional interventions (2011)
3. Evans, J.S., Barston, J.L., Pollard, P.: On the conflict between logic and belief in syllogistic reasoning. Mem. Cogn. **11**(3), 295–306 (1983)
4. Lynch, C., Ashley, K., Aleven, V., Pinkwart, N.: Defining ill-defined domains; a literature survey. In: Proceedings of the Workshop on Intelligent Tutoring Systems for Ill-defined Domains at the 8th International Conference on Intelligent Tutoring Systems, pp. 1–10 (2006)
5. Silverman, C., Singer-Vine, J.: Most Americans who see fake news believe it, new survey says. BuzzFeed News (2016). https://www.buzzfeed.com/craigsilverman/fake-newssurvey