

# An Argument-Based Approach to Cope with Trust and Pluralism in Web News Reports

Fernando M. Sagui<sup>†§</sup> Ana G. Maguitman<sup>†§</sup>  
Carlos I. Chesñevar<sup>†§</sup> Guillermo R. Simari<sup>†</sup>  
{fms, agm, cic, grs}@cs.uns.edu.ar

<sup>†</sup>Laboratorio de Investigación y Desarrollo en Inteligencia Artificial  
Departamento de Ciencias e Ingeniería de la Computación

Universidad Nacional del Sur, Av. Alem 1253, (8000) Bahía Blanca, Argentina\*

<sup>§</sup>CONICET - Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

## Abstract

Due to the huge amount of multi-source news that are available on the Web at any time, it is crucial to provide intelligent mechanisms to select and rank news reports. Over the last few years, a number of approaches based on criteria such as freshness, relevance and viewer profile have been proposed. However, most existing news processing services do not deal with credibility as a subjective notion from a qualitative perspective, and do not provide mechanisms to cope with controversial news reports. To fill this gap, this paper proposes a news service framework that brings the notions of trust and pluralism into play. The proposed framework is based on a set of basic principles characterizing the nature of trust. In our proposal, trust is modeled using Defeasible Logic Programming, a general-purpose defeasible argumentation formalism based on logic programming. Our approach helps identify antagonism among sources of news and facilitates the analysis of opposing positions. This allows us to integrate dialectical reasoning into a news recommender system, which has the capability of providing a reasoned basis for the news presented to the viewer.

**Keywords:** newsfeeds, personalized news, web, argumentation, trust, credibility, trustworthiness, pluralism

## 1 INTRODUCTION

According to several studies by Nielsen/NetRatings [4], reading news has become one of the most important activities on the Web. The number of visitors to news websites has steadily increased over the last years, and the abundant supply of online news is a clear indication of users' urge to be informed. Moreover, the existence of multiple sources of news provides an opportunity to access pluralistic opinions, which can be regularly found on the media.

Nowadays, there are several commercial multi-source news providers on the web, such as Google News [2], Yahoo! News [3], MSNBC [1], etc. Although none of them has disclosed the technical details underlying the way news are selected, aggregated and ranked, it is evident that factors such as freshness, sources and popularity are taken into account. The information

---

\*Partially supported by CONICET (PIP 5050), Agencia Nacional de Promoción Científica y Tecnológica (PICT 2005, Nro. 32373) and Universidad Nacional del Sur (24/N016 and 2/ZN13).

provided in news reports may not always be fully verifiable and therefore another important factor that can help select news is *trust* or *credibility*.

Research on multi-source news has generally overlooked the dynamics of news credibility, or if considered, it has been studied through quantitative approaches (e.g. [13]). There is also documented evidence [7] of Google News’s plans to build a database of news source credibility based on information such as average story length, number of staff a news source employs, the volume of internet traffic to its website and the number of countries accessing the site. Google approach to dealing with news credibility is to take all these and other parameters to create a single value used to rank the results of any news search. However, a foolproof approach to deal with news trust has not been developed yet and multi-source news services remain vulnerable to credibility breaches.<sup>1</sup>

A purely quantitative perspective to news credibility has several limitations. On the one hand, the absence of a formal model underlying quantitative approaches makes it hard to provide viewers with a justification of why certain news should be trusted. Because quantitative approaches are not equipped with inference capabilities, much of the implicit information remains undiscovered. On the other hand, they are incapable of dealing formally with the defeasible nature of trust. In addition, because trust is to a great degree subjective, quantifying trust by combining measures coming from a pool of credibility assessments may not be entirely realistic.

The goal of this research paper is to define a qualitative and personalized trust-based news service. The service will allow news viewers to access and compare the trustworthiness of news sources and their reports. Viewers’ trust statements on sources and reports can be based on the viewers’ subjective beliefs or, when absent, trust assumptions can be obtained indirectly from other viewers’ beliefs. However, in order to derive trust from other viewers, a trust relationship between viewers must exist.

In this proposal, trust is modeled using DeLP, a defeasible argumentation framework based on logic programming [?]. This allows us to integrate dialectical reasoning into a news service, which will provide a reasoned basis for the news presented to the user.

The paper is organized as follows. The next section presents the background and motivations for our proposal, reviewing DeLP and discussing the problem of ranking and trusting Web news. Section 3 proposes a set of postulates for news trust and shows how to represent them by means of a set of DeLP rules. In section 4 we illustrate the proposal with a an example. Finally, section 5 overviews related work and section 6 outlines our conclusions.

## 2 BACKGROUND AND MOTIVATIONS

### 2.1 Defeasible logic programming

Defeasible logic programming (DeLP) [9] is a general-purpose defeasible argumentation formalism based on logic programming, intended to model inconsistent and potentially contradictory knowledge. A defeasible logic program has the form  $\mathcal{P} = (\Pi, \Delta)$ , where  $\Pi$  and  $\Delta$  stand for *strict* and *defeasible* knowledge, respectively. The set  $\Pi$  involves *strict rules* of the form  $P \leftarrow Q_1, \dots, Q_k$  and *facts* (strict rules with empty body), and it is assumed to be *non-contradictory* (i.e., no complementary literals  $P$  and  $\sim P$  can be inferred, where  $\sim P$  denotes the contrary of  $P$ ). The set  $\Delta$  involves *defeasible rules* of the form  $P \rhd Q_1, \dots, Q_k$ , which

---

<sup>1</sup>A famed example of this vulnerability problem is that of a fifteen-year old teenager who put out a fake release saying he had been hired by Google and within hours, his release was picked up in Google News.

stand for “ $Q_1, \dots, Q_k$  provide a tentative reason to believe  $P$ .” Rules in DeLP are defined in terms of *literals*. A literal is an atom  $A$  or the strict negation ( $\sim A$ ) of an atom. Default negation (denoted *not*  $A$ ) is also allowed in the body of defeasible rules (see [9] for details).

Deriving literals in DeLP results in the construction of *arguments*. An argument  $\mathcal{A}$  for a literal  $Q$  (denoted  $\langle \mathcal{A}, Q \rangle$ ) is a (possibly empty) set of ground defeasible rules that together with the set  $\Pi$  provide a proof for a given literal  $Q$ , satisfying the additional constraints of *non-contradiction* (i.e., an argument should not allow the derivation of contradictory literals) and *minimality* (i.e., the set of defeasible information used to derive  $Q$  should be minimal). Note that arguments are obtained by a mechanism similar to the usual query-driven SLD derivation from logic programming, performed by backward chaining on *both* strict and defeasible rules; in this context a negated literal  $\sim P$  is treated just as a new predicate name *no\_P*. In DeLP, arguments provide tentative support for claims (literals). Clearly, as a program  $\mathcal{P}$  represents incomplete and tentative information, an argument  $\langle \mathcal{A}, Q \rangle$  may be *attacked* by other arguments also derivable from  $\mathcal{P}$ . An argument  $\langle \mathcal{B}, R \rangle$  is a *counter-argument* for  $\langle \mathcal{A}, Q \rangle$  whenever a subargument  $\langle \mathcal{A}', Q' \rangle$  (with  $\mathcal{A}' \subseteq \mathcal{A}$ ) in  $\langle \mathcal{A}, Q \rangle$  can be identified, such that  $\langle \mathcal{B}, R \rangle$  and  $\langle \mathcal{A}', Q' \rangle$  cannot be simultaneously accepted since their joint acceptance would allow contradictory conclusions to be inferred from  $\Pi \cup \mathcal{A}' \cup \mathcal{B}$ . If the attacking argument  $\langle \mathcal{B}, R \rangle$  is preferred over  $\langle \mathcal{A}', Q' \rangle$ , then  $\langle \mathcal{B}, R \rangle$  is called a *defeater* for  $\langle \mathcal{A}, Q \rangle$ . The preference criterion commonly used is *specificity* [9], preferring those arguments which are more direct or more informed, although other criteria could be adopted.

In DeLP the search for defeaters for a given argument  $\langle \mathcal{A}, Q \rangle$  prompts a recursive process, resulting in the generation of a *dialectical tree*: the root node of this tree is the original argument at issue, and every children node in the tree is a defeater for its parent. Additional restrictions help to avoid circular situations when computing branches in a dialectical tree, guaranteeing that every dialectical tree is finite (see [9] for details). Nodes in the tree can be marked either as *defeated* (*D*-nodes) or as *undefeated* (*U*-nodes). The marking of the dialectical tree is performed as in an AND-OR trees: leaves are always marked as undefeated nodes (as they have no defeaters); inner nodes can be marked either as undefeated (if and only if every of its children nodes is marked as defeated) or as defeated (whenever at least one of its children has been marked as undefeated). The original argument  $\langle \mathcal{A}, Q \rangle$  (the root of tree) is deemed as ultimately acceptable or *warranted* whenever it turns out to be marked as undefeated after applying the above process.

Figure 1 shows an example of how DeLP can be used to represent commonsense knowledge about spiders. In this sample program there are four defeasible rules (e.g. spiders are usually dangerous, spiders which are dead are usually not dangerous, etc.). The program also includes some facts about a particular spider in a given situation (e.g. a black widow spider which looks dead, but moves when touched). By performing the query `dangerous(black_widow)`, DeLP allows us to conclude that we have a warranted argument supporting the claim that this spider is dangerous. Note that this involves computing a dialectical tree, which in this particular example involves just a single branch. As the query `dangerous(black_widow)` is supported by a warranted argument, the answer provided by DeLP is *yes*.

Note also that the computation of the dialectical tree is performed automatically by the DeLP interpreter on the basis of the program available. This process is based on an abstract machine which extends Warren’s abstract machine for PROLOG [9]. Given a DeLP program  $\mathcal{P}$ , solving a query  $Q$  with respect to  $\mathcal{P}$  may result in four possible answers: YES (there is at least one warranted argument  $\mathcal{A}$  for  $Q$ ); NO (there is at least one warranted argument  $\mathcal{A}$  for  $\sim Q$ ); UNDECIDED (none of the previous cases hold); and UNKNOWN ( $Q$  is not present in the

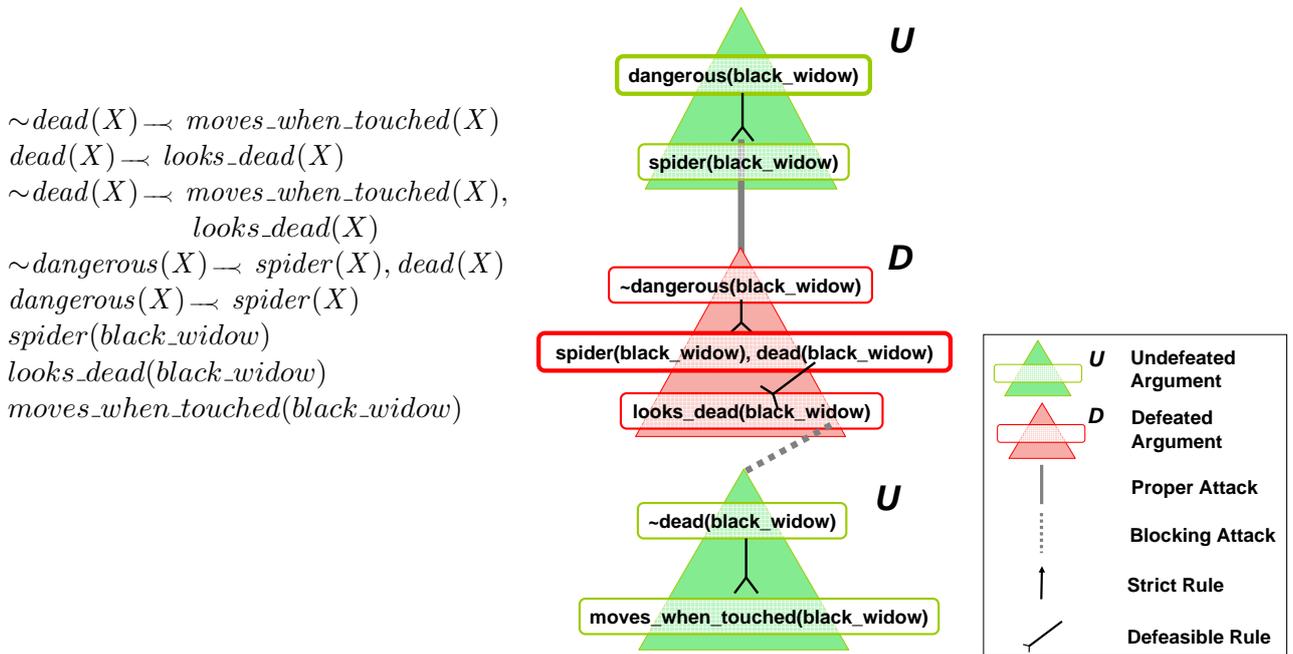


Figure 1: A sample DeLP program about spiders, and the dialectical analysis associated with solving the query `dangerous(black_widow)`

program signature). The emerging semantics is skeptical, computed by DeLP on the basis of the goal-directed construction and marking of dialectical trees, which is performed in a depth-first fashion. Additional facilities (such as visualization of dialectical trees, zoom-in/zoom-out view of arguments, etc.) are integrated in the DeLP environment to facilitate user interaction when solving queries.

## 2.2 Ranking and Trusting News

The problem of ranking Web news has attracted much attention in recent years. There are several reasons why measures of page authority such as PageRank [5] cannot be directly applied at the moment of ranking Web news. Differently from what happens with webpages, the Internet newspapers rarely use linking. Moreover, breaking news usually have priority over previous news because viewers prefer to see information about news events as soon as they take place. However, fresh news usually have very few incoming links, which precludes the application of link analysis algorithms to favor fresh news over stale ones. A ranking model that gives high priority to fresh news, however, will have some deficiencies. Reports on fresh news tend to be incomplete and many stories presented as breaking news are revised when additional information becomes available.

Another approach to rank news could be based on news popularity, estimated by monitoring the number of viewers accessing a report or by a system of voting on favorite stories. However, news popularity may not reflect the real value a news has for individual viewers.

Usually a viewers has to decide whether a news report is worth reading and whether the facts described in the report are credible. Unfortunately, developing an algorithm for Web news selection and ranking is very difficult because it needs to combine many, sometimes conflicting, aspects. The level of trust a viewer has on a piece of news is not necessarily associated with measures of news authority or popularity, and it may even be negatively correlated with news

freshness. A news service that uses a trust management system can support the viewer in making the decision by selecting reports from trusted sources or based on another trustworthy viewer's opinion.

There are important aspects of trust that need to be considered in order to develop a realistic modeling of trust, as well as reliable and usable services based on this notion:

- **Trust should be Justified.** Most existing news ranking services act as “black boxes”, because they refuse to disclose how they select certain news or rank them in certain ways. This results in trustworthiness issues because they do not provide viewers with a justification of why certain news should be trusted.
- **Trust is Defeasible.** News reports trusted by some viewer can be superseded by other reports carrying more authority, say from CNN or some other trusted source. In the meantime, news agencies are subject to time constraints, which results in the publication of reports with incomplete information or inaccurate information. Trust on such reports could be revoked by the release of other more recent ones.
- **Trust is Subjective.** Like many aspects of the Web, news is becoming a collaborative activity. However, judgment of news credibility is idiosyncratic. Therefore, models that deal with trust as an objective notion are unrealistic. In this case, the opinion from the “wisdom of the crowds” may not be as useful as the viewer's personal opinion, or the opinion of another trusted viewer.

The rest of the paper is aimed at providing a framework for dealing with news credibility on the Web, in which the above characteristics of trust are taken into consideration.

### 3 A FORMAL FRAMEWORK FOR REASONING ABOUT NEWS TRUST

A system that reasons about news trust should take a number of ingredients into consideration:

- **Reports.** A report or news article is a written communication of a news event prepared by a specific news agency (source). When a report is made available on the Web through an newsfeed, we can identify fields such as *title*, *source*, *timestamp*, *description*, *category* and *link to news content*. Other information related to the report such as *author* can also be derived on certain situations.
- **Sources.** The source of a news article is the agency in charge of supplying the report to be used by the media.
- **Viewers.** A viewer is a user of the news service. The system maintains a pool of viewers. Viewers can provide trust statements about reports, sources and other viewers.
- **Trust/Distrust Statements.** A trust (distrust) statement is an explicit assertion of the fact that a viewer trust (distrust) a report, a source or another viewer. These statements allow to infer implicit trust relations, which are useful to provide recommendations to the viewer based on trust.

### 3.1 Postulates for Trust Statements

This subsection presents a set of postulate for trust (and distrust) statements. The set of postulates will embody the general intuitions about the way trust and distrust statements could be derived from existing ones.

**Postulate 1.** A report coming from a trusted source will typically be trusted.

**Postulate 2.** A report coming from a distrusted source will typically be distrusted.

**Postulate 3.** A report trusted by a trusted viewer will typically be trusted.

**Postulate 4.** A report distrusted by a trusted viewer will typically be distrusted.

**Postulate 5.** A source trusted by a trusted viewer will typically be trusted.

**Postulate 6.** A source distrusted by a trusted viewer will typically be distrusted.

An interesting situation will arise when two conflicting conclusions can be reached. For example, a trusted viewer distrust a report, but the report was released by a trusted source. Assuming that in general we prefer to base our opinion on information provided by trusted sources, we can add the following two postulates:

**Postulate 7.** A report coming from a trusted source will typically be trusted, even if it is distrusted by a trusted viewer.

**Postulate 8.** A report coming from a distrusted source will typically be distrusted, even if it is trusted by a trusted viewer.

Additional postulates could be added. For example, if some viewer has a very good reputation for fact-checking, we will prefer to trust this viewer's opinion even if it conflicts with other viewers' opinions. Other postulates that could be added to the list may include references to the timestamp of the report. For instance, a more recent report will be trusted over an outdated one, unless the report is just out (Due to the eagerness to publish the story some fresh news reports may not be as reliable as older ones.) The list of postulates could be extended indefinitely, including references to news author, country of origin of the source, news category, etc. Some postulates could be personalized, because different viewers may disagree on the conclusion that should be adopted given certain facts. For the sake of simplicity, we will take postulates 1 to 8 as the core postulates for our trust-management system.

### 3.2 Using DeLP to Reason about News Trust and Pluralism

[[REVISAR Y EXTENDER]] Postulates 1 to 8 can be naturally modeled using the following DeLP rules:

$$\begin{aligned}
 \textit{trust\_report}(V, R) &\multimap \textit{report\_source}(R, S), \textit{trust\_source}(V, S) & (R_1) \\
 \sim\textit{trust\_report}(V, R) &\multimap \textit{report\_source}(R, S), \sim\textit{trust\_source}(V, S) & (R_2) \\
 \textit{trust\_report}(V, R) &\multimap \textit{trust\_viewer}(V, V1), \textit{trust\_report}(V1, R) & (R_3) \\
 \sim\textit{trust\_report}(V, R) &\multimap \textit{trust\_viewer}(V, V1), \sim\textit{trust\_report}(V1, R) & (R_4) \\
 \textit{trust\_source}(V, S) &\multimap \textit{trust\_viewer}(V, V1), \textit{trust\_source}(V1, S) & (R_5) \\
 \sim\textit{trust\_source}(V, S) &\multimap \textit{trust\_viewer}(V, V1), \sim\textit{trust\_source}(V1, S) & (R_6) \\
 \textit{trust\_report}(V, R) &\multimap \textit{report\_source}(R, S), \textit{trust\_source}(V, S), \\
 &\quad \textit{trust\_viewer}(V, V1), \sim\textit{trust\_report}(V1, R) & (R_7) \\
 \sim\textit{trust\_report}(V, R) &\multimap \textit{report\_source}(R, S), \sim\textit{trust\_source}(V, S), \\
 &\quad \textit{trust\_viewer}(V, V1), \textit{trust\_report}(V1, R) & (R_8)
 \end{aligned}$$

## 4 A WORKED EXAMPLE

[[REVISAR Y EXTENDER]]

|  |            |
|--|------------|
| $report\_source(true\_news, the\_truthteller)$ | $(F_1)$    |
| $report\_source(false\_news, the\_corker)$     | $(F_2)$    |
| $report\_source(some\_news, the\_incog)$       | $(F_3)$    |
| $trust\_source(joe, the\_truthteller)$         | $(F_4)$    |
| $\sim trust\_source(tom, the\_corker)$         | $(F_5)$    |
| $\sim trust\_report(ann, true\_news)$          | $(F_6)$    |
| $trust\_report(ann, false\_news)$              | $(F_7)$    |
| $\sim trust\_report(ann, some\_news)$          | $(F_8)$    |
| $trust\_report(tom, some\_news)$               | $(F_9)$    |
| $trust\_viewer(joe, ann)$                      | $(F_{10})$ |
| $trust\_viewer(joe, tom)$                      | $(F_{11})$ |

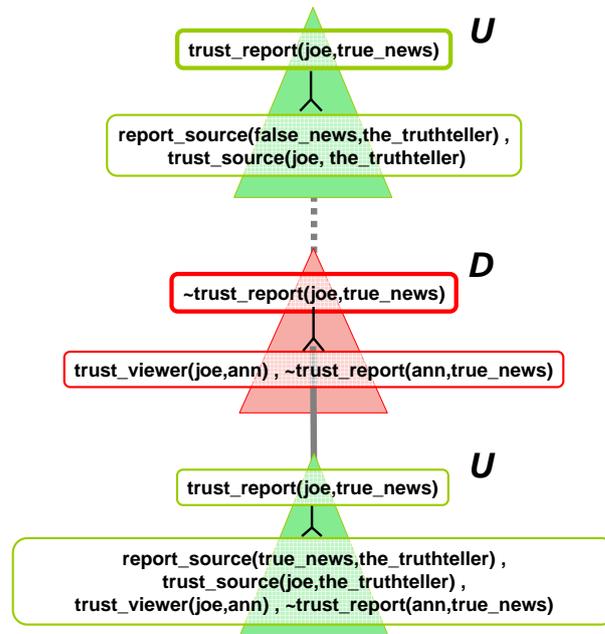


Figure 2: Yes

## 5 RELATED WORK

[[REVISAR Y EXTENDER]]

A variety of methods have been proposed and a number of systems have been developed to facilitate access to news on the web. NewsInEssence [15] is a system that searches and clusters related news. QCS [6] is a tool that facilitates the task of grouping and categorizing news. In [12] a method is proposed to search web articles while TV news are on the air. Other tools [16] automatically extracts news from websites. Velthune [11] is a search engine that extracts information both from the web and from newsfeeds. Another system, NewsJunkie [8],

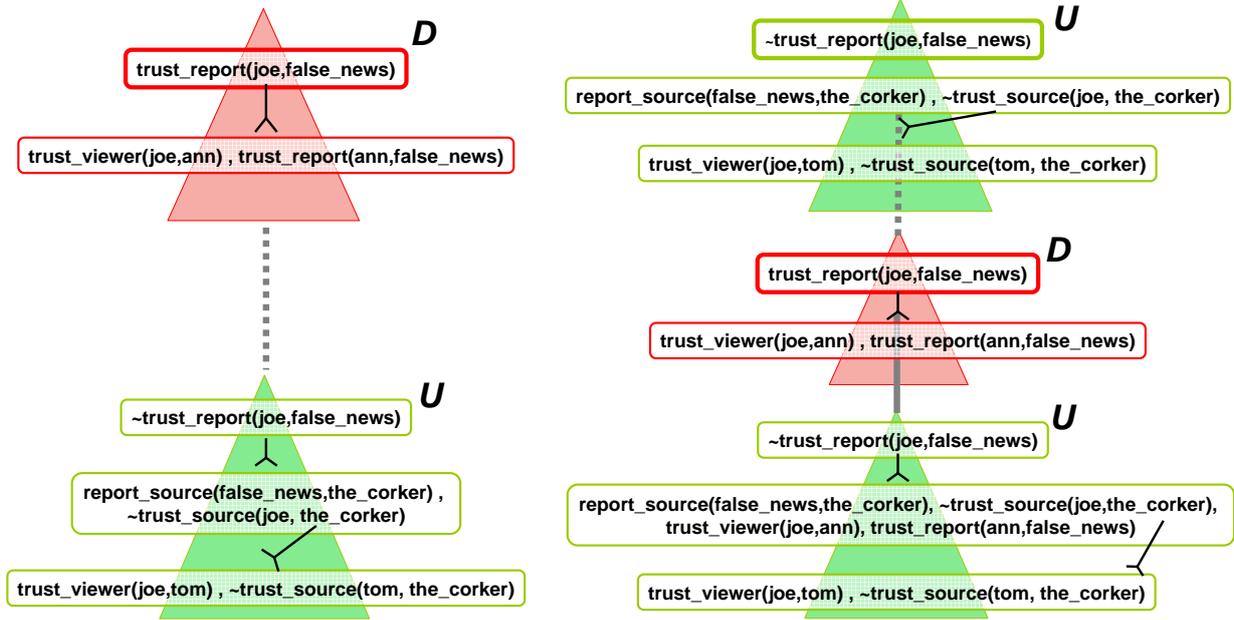


Figure 3: No

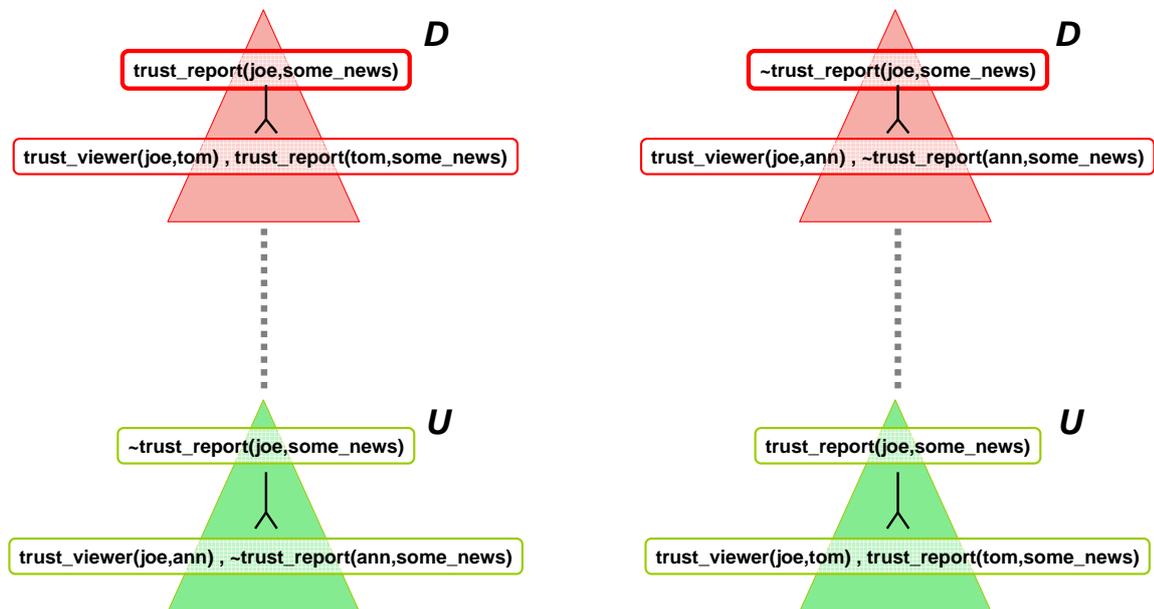


Figure 4: Undecided

has the purpose of identifying novel news in the context of stories previously reviewed by the user.

PolyNews [14]

## 6 CONCLUSIONS

[[REVISAR Y EXTENDER]]

Contribution: formal framework to deal with news: notions have been identified and formally defined.

As a first step we have implemented a multi-source news service that monitors several newsfeeds and populates a database of news ([http://lidia.cs.uns.edu.ar/delp\\_client](http://lidia.cs.uns.edu.ar/delp_client)). We are currently developing algorithms for clustering news reports based on the news event the reports describe. In the near future, we expect to integrate this environment with information about viewers and their trust statements, as well as to equip the system with reasoning capabilities based on DeLP. The DeLP environment is available online to test at [http://lidia.cs.uns.edu.ar/delp\\_client](http://lidia.cs.uns.edu.ar/delp_client).

A user is much more likely to believe statements from a trusted acquaintance than from a stranger. Therefore, trust could be quantified and trust values could be transitively obtained from other viewer and scaled down depending on the viewer providing the information. In this sense we expect to study ways of extending our model of trust by adding (... decir algo sobre P-DeLP o algún otro método que nos sirva para modelar este aspecto cuantitativo ...). An example of trust and distrust propagation method can be found in [10]

An argument-based approach allows the exploration of arguments that challenge the viewer perspective. Using an argument-based framework can help the viewer review previously maintained trust statement, as it is possible to analyze the justification in favor and against trusting some report. This could also help discover events in which biased sources present facts in a convincing manner, but neglect to include other important facts that are against their position. Because conclusions are justified by the system, the viewer is able to go back and figure out where errors were revealed. But in practice many facts remain unverifiable and there is room for ideology and interpretation. The proposed framework is faithful to this aspect of reality.

## REFERENCES

- [1] <http://newsbot.msnbc.msn.com/>.
- [2] <http://news.google.com/>.
- [3] <http://news.yahoo.com/>.
- [4] <http://www.nielsen-netratings.com/>.
- [5] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107-117, 1998.
- [6] Daniel Dunlavy Ddunlavy. QCS: A tool for querying, clustering, and summarizing documents.
- [7] Barry Fox. Google searches for quality not quantity. *New Scientist magazine*, 2497:24. 30 April 2005.
- [8] E. Gabrilovich, S. Dumais, and E. Horvitz. Newsjunkie: Providing personalized newsfeeds via analysis of information novelty, 2004.

- [9] A. García and G. Simari. Defeasible Logic Programming: An Argumentative Approach. *Theory and Practice of Logic Programming*, 4(1):95–138, 2004.
- [10] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *International World Wide Web Conference (WWW2004)*, 2004.
- [11] A. Gulli. The anatomy of a news search engine. In *WWW 2005*, Chiba, Japan, May 2005.
- [12] M. Henzinger, B. Chang, B. Milch, and S. Brin. Query-free news search, 2003.
- [13] Ryosuke Nagura, Yohei Seki, Noriko Kando, and Masaki Aono. A method of rating the credibility of news documents on the web. In *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 683–684, New York, NY, USA, 2006. ACM Press.
- [14] Souneil Park, Seungwoo Kang, and Junehwa Song. Polynews: Delivering multiple aspects of news to mitigate media bias. Technical report, KAIST, 2006.
- [15] Dragomir R. Radev, Sasha Blair-Goldensohn, Zhu Zhang, and Revathi Sundara Raghavan. NewsInEssence: A system for domain-independent, real-time news clustering and multi-document summarization.
- [16] D. Reis, P. Golgher, A. Silva, and A. Laender. Automatic web news extraction using tree edit distance, 2004.