

FORECAST2023: A Forecast and Reasoning Corpus of Argumentation Structures

Kamila Górska, John Lawrence, Chris Reed

Centre for Argument Technology, University of Dundee, UK

Forecasting Competitions

Forecasting Competitions explore the ability to make accurate predictions in decision-making which can be applied to many domains[1]. We use data from the Hybrid Forecasting Competition[2], which includes the forecaster's rationales, providing an insight into their thought process which we model to provide insight into the argumentative strategies used.

Example of a Forecast

Will France's President Emmanuel Macron experience a significant leadership disruption between 3 April 2019 and 29 November 2019?

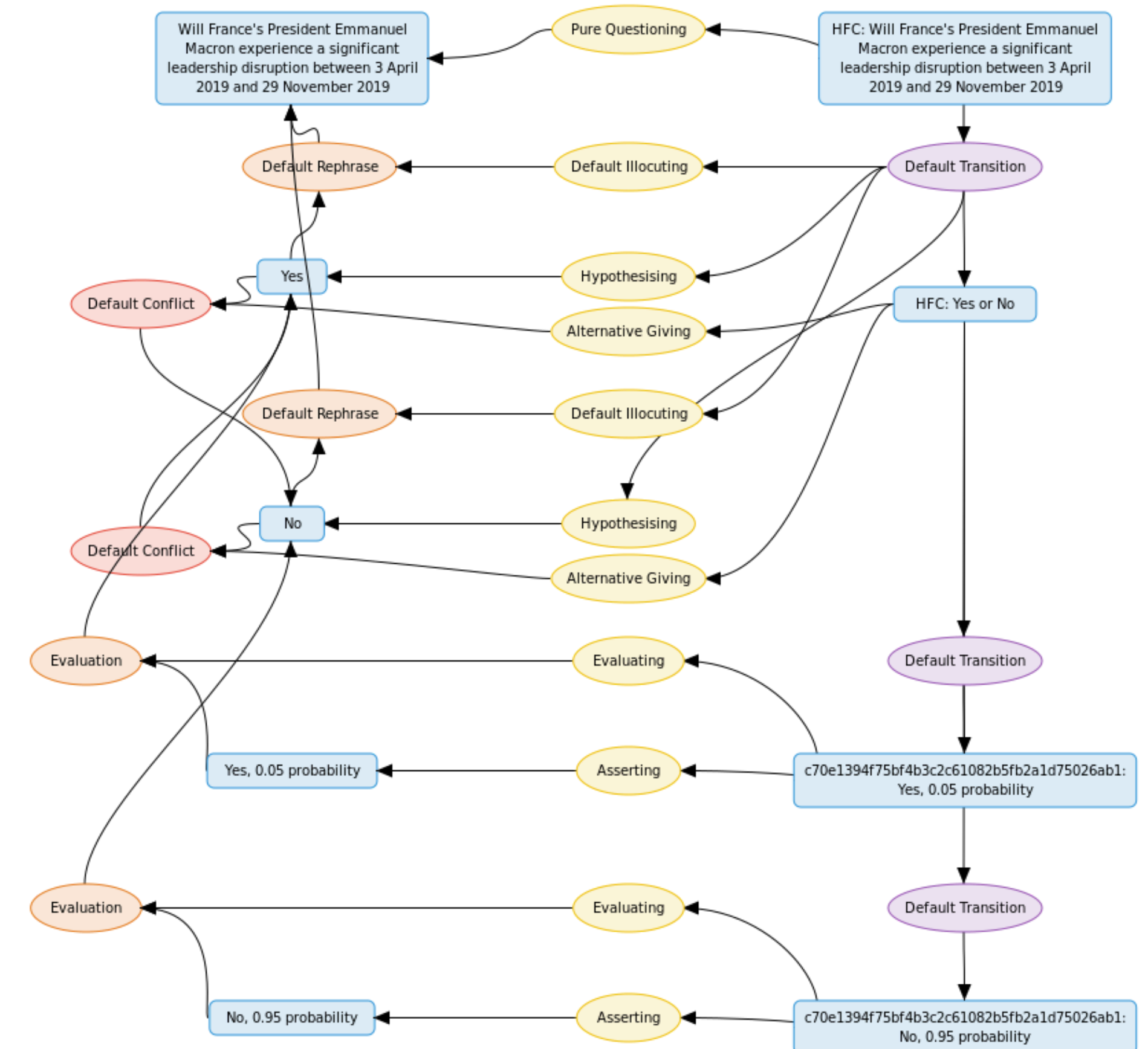
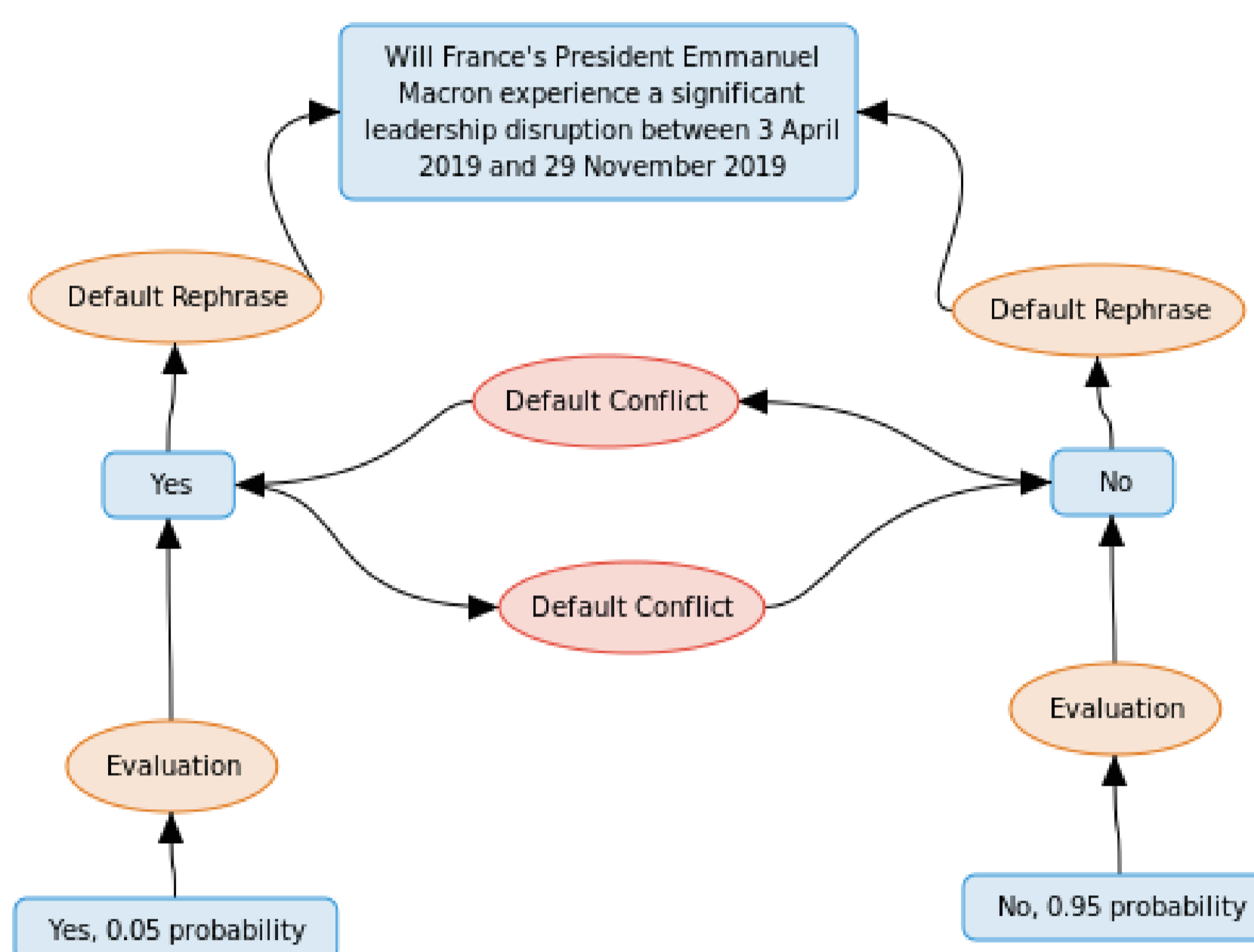
Yes: 0.05
No: 0.95

Rationale

I believe with the yellow vest protests put on by the working class who have been hit with low wages and heightened gasoline taxes are a sign of the people's abilities, but I do not believe this is something that will take Macron out of office. Macron has attempted to meet voters demands and is trying to rectify the situation by being more involved and less out of touch. Macron is hosting or involved in political discussions, debates and meetings to address the unrest in his country. I do think he's made plenty of mistakes that caused a lot of damage (albeit some being done by the protestors), but I believe he is regaining 'popularity' enough to maintain office.

Automated Processing

A sample of 120 forecasts was chosen to be annotated in Inference Anchoring Theory (IAT), which is used for the annotation of the largest argumentation datasets[3]. IAT models the argument structure, the speech acts, and the dialogue rules. The first step of annotation is completed automatically to speed up the process, by creating IAT annotation of the Question, the possible answers, and the forecasted probabilities of each answer occurring.



- The Question is annotated as an instance of 'Pure Questioning' which is an information-seeking question.
- The potential answers provided by the competition are annotated as acts of 'Hypothesising', and are alternatives to each other which is reflected through the conflict between them.
- The forecasts made by the forecasters are assertions which evaluate the corresponding hypothesis.

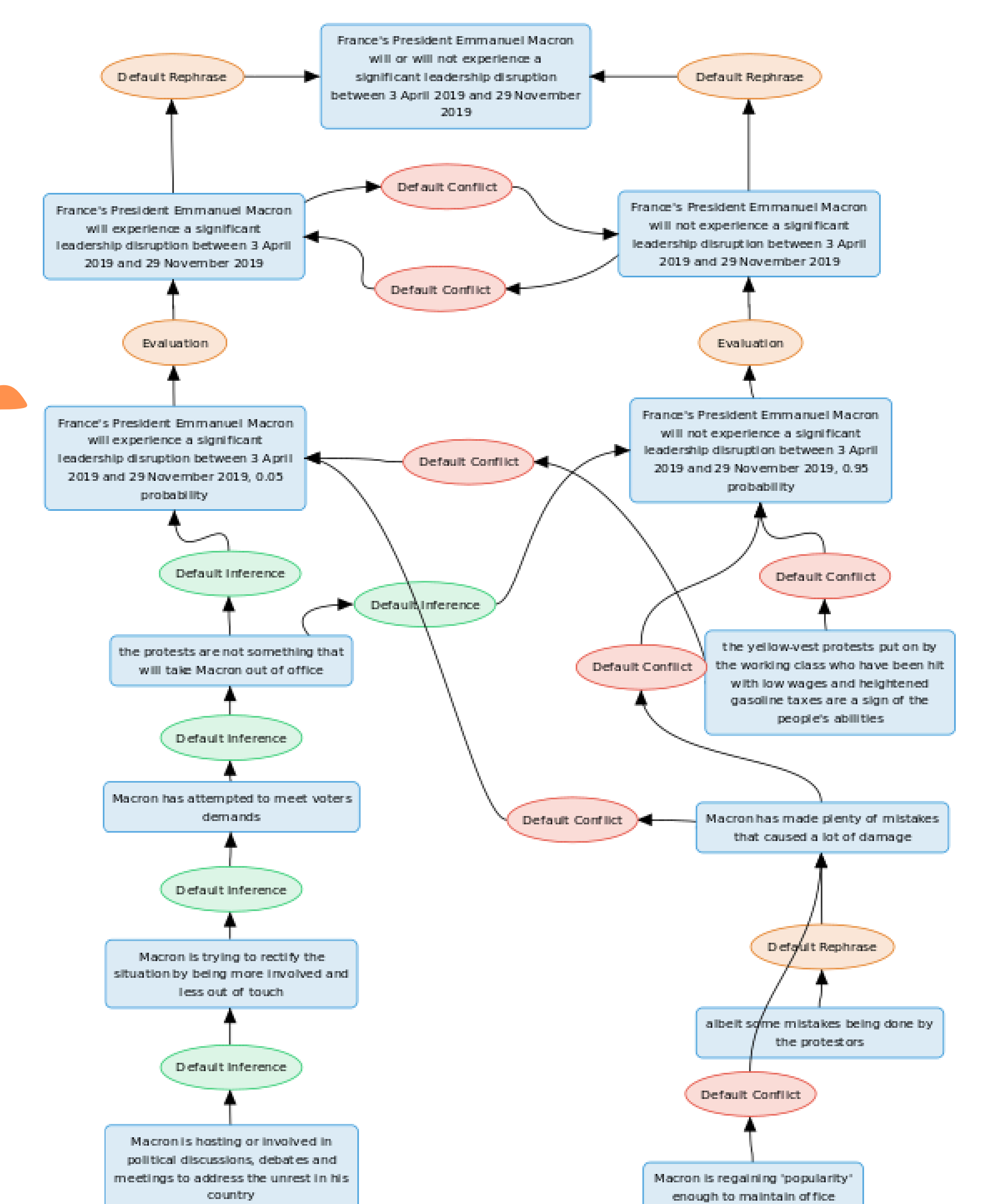
FORECAST2023 Dataset

Feature	Count
Rationales	205
-- (which are updates)	85
Questions	100
Hypotheses	353
Support Relations	968
Conflict Relations	946
-- (which are not alternative hypotheses)	74
Words	31556

The work presented here covers the creation of a unique dataset of prediction rationales that provides a uniquely fine-grained and close characterisation of the structure of argumentation, with potential impact on forecasting domains from intelligence analysis to investment decision-making.

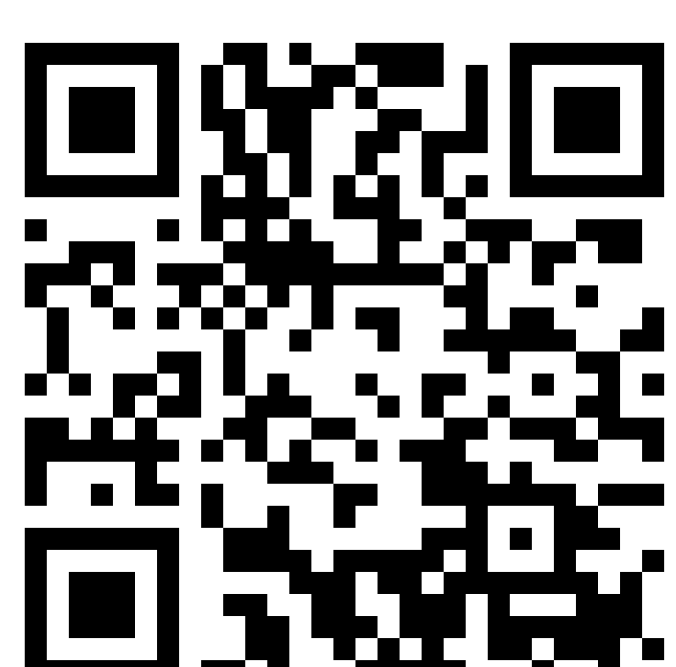
The table on the left shows the counts of various features of the final annotated corpus. Out of the 205 rationales, 85 were updates to previous forecasts making up 41.5% of the corpus. The quality and length of the writing do not differ drastically between the rationales, however, the rationales do differ in the objectivity of the forecaster and focus on arguing for one hypothesis as opposed to having arguments to support or attack each possible option.

Manual Annotation



The rationale is annotated manually by a team of 11 analysts who complete the IAT analysis and connect the rationale to the forecasts above through argument relations. Rationales provide supports and attacks of the forecasts, and reasoning throughout the rationale itself.

Find Out More



FORECAST2023 provides a first step towards the argumentative analysis of forecasting rationales, which opens the door to deeper analysis among other things: how argumentative strategies affect decision-making, the key strategies accurate forecasters use to come to decisions, and exploring how decision-making can be improved across various domains. The fully annotated corpus, the code used to develop the automated annotation and pre-processing, and the manual annotation guidelines are all openly available online.

- 1: C. W. Karvetski, et al. 2022. What do forecasting rationales reveal about thinking patterns of top geopolitical forecasters? International Journal of Forecasting, 38(2):688-704.
- 2: D. M Benjamin, et al. 2023. Hybrid forecasting of geopolitical events. AI Magazine. 44(1), 112-128
- 3: A. Hautli-Janisz, et al. QT30: A corpus of argument and conflict in broadcast debate. In Proceedings of the 13th Language Resources and Evaluation Conference, pages 3291-3300. European Language Resources Association (ELRA).