1 Analytics List

1.1 Rationale-level

- **Brier Forecast Accuracy**: Brier scoring is used to calculate the squared error of the prediction. Each forecast receives a score between 0 and 1, with 0 being the best (no difference between the prediction and outcome), and 1 being the worst (100% difference between prediction and outcome).

$$\text{Brier Score} = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

  Where N is the number of predictions, p_i is the predicted probability of the event occurring for the i-th observation, o_i is the actual outcome for the i-th observation ($o_i \in \{0, 1\}$).
- **Event Outcomes**: The binary outcome metric measures whether the outcome the forecaster believed to be most likely, was the outcome which came true, discounting the confidence of the forecaster. The highest probability assigned by the forecaster was taken as the outcome they believed to be most likely, if this is the event which did occur, the forecast is scored as successful. If the option with the highest assigned probability was not the actual outcome, or there was no highest assigned option, the forecast is scored as unsuccessful.

1.2 Map(/hypothesis)-level

A forecast and any hypotheses within a forecast are analysed as argument graphs. For the identification of hypothesis subgraphs within the forecast graph, see the description in the main paper.

In characterising types of argumentative relations as nodes, an RA-node represents a relation of support and a CA-node a relation of conflict. An RA-node is used as an intermediary to express support by nodes with edges towards the RA (premises) for nodes with edges from the RA (conclusions), and equivalently a CA-node is used to express an attack by nodes with edges towards the CA on nodes with edges from the CA.

Almost all analytics in this section are applied both at map level and at hypothesis level. The exceptions are *RA word density*, *CA word density* and *Argument word density* as these use the wordcount of the entire original text, rather text segments explicitly associated with nodes of the graph. Appropriate association of unassigned surrounding text with one particular hypothesis (or none) is non-trivial, and consequently these analytics do not generalise without adjustment for use on hypothesis subgraphs.

We define arguments and paths in arguments as follows.

Let an argument a be a tuple $\langle N, E\rangle$ of a set of nodes $N = \{n_1, \dots, n_n\}$ and edges $E = \{\langle n_i, n_j\rangle, \dots\}$ with the following properties:

- $\forall n \in N : n \in \text{I} \cup \text{RA} \cup CA$
  Nodes in the argument are either information nodes ($n \in$ I), nodes of application of schemes of support ($n \in$ RA), or nodes of application of schemes of conflict ($n \in$ CA).
- $\forall \langle n_i, n_j\rangle \in E : n_i, n_j \in N$ and $n_i \neq n_j$
  Edges connect distinct nodes in the graph.
- $\nexists \langle n_i, n_j\rangle \in E$ s.t. $n_i, n_j \in$ I
  No edge directly connects two I-nodes.
- $\forall \langle n_i, n_j\rangle \in E$: if $n_i \in \text{RA} \cup \text{CA}, \exists \langle n_a, n_i\rangle \in E$; if $n_j \in \text{RA} \cup \text{CA}, \exists \langle n_j, n_z\rangle \in E$
  RA and CA nodes must have an incoming and outgoing edge.

For an argument a, a path p in the argument is a set $p \subseteq E$ such that:

- $|\{\langle n_i, n_j\rangle \in E : \nexists n_k : \langle n_k, n_i\rangle \in E\}| \leq 1$
  At most one edge begins with a node which is not itself the end-point of another edge in the path.
- $|\{\langle n_i, n_j\rangle \in E : \nexists n_k : \langle n_j, n_k\rangle \in E\}| \leq 1$
  At most one edge ends with a node which is not itself the start-point of another edge in the path.

Therefore, all other edges in the path begin with a node which is itself the end-point of another edge in the path, and end with a node which is itself the start-point of another edge in the graph.

Let the expanded argument graph which includes discourse information be a^d, a tuple $\langle N^d, E^d\rangle$ such that $N \subseteq N^d$ and $E \subseteq E^d$. Additionally, let a subset of information L nodes be distinguished as locution nodes L $\subset$ I, present in the expanded argument graph only, $N \cap L = \emptyset$. The expanded graph a^d holds the following properties:

- $\forall n \in N^d : n \in N \cup \text{TA} \cup \text{YA}$
  The expanded graph additionally contains nodes which express a discourse transition ($n \in$ TA) or use of illocutionary force ($n \in$ YA).
- $\forall \langle n_i, n_j\rangle \in E^d : n_i, n_j \in N^d$ and $n_i \neq n_j$
  Edges connect distinct nodes in the graph.

Let t_{n_i} indicate the text associated with a node n_i.

Finally, the provision of exclusive alternatives, e.g. choices in a question which a forecaster may choose between, is expressed through the presence of a conflict (CA) between the options, connected to the discourse via the speech act of *Alternative Giving*. To avoid particular inflation of the count of CAs due to the number of potential answers available for a given forecast, these option-expressing CAs explicitly excluded for analytics involving CA nodes.

$$\begin{aligned} N^{ex} = N - \{n_i \in N : n_i \in \text{CA},\\ \exists \langle n_j, n_i\rangle \in E^d,\\ t_{n_j} = \text{`Alternative Giving'}\} \end{aligned}$$

N^{ex} is the subset of nodes N excluding any CA nodes with an incoming edge from an *Alternative Giving* node. An expanded graph which both includes discourse and excludes *Alternative Giving*-connected CA nodes can be denoted as $N^{d,ex}$.

Simple analytics

- **Support count** Number of RA-nodes in the graph.

$$|\{n \in N : n \in \text{RA}\}|$$

- **Conflict count** Number of CA-nodes in the graph.

$$|\{n \in N^{ex} : n \in \text{CA}\}|$$

- **Support word density** RA count relative to wordcount of the original text.

$$\frac{|\{n \in N : n \in \text{RA}\}|}{\textit{wordcount}}$$

- **Conflict word density** CA count relative to wordcount of the original text.

$$\frac{|\{n \in N^{ex} : n \in \text{CA}\}|}{\textit{wordcount}}$$

- **Argument word density** Combined RA and CA count relative to wordcount of the original text.

$$\frac{|\{n \in N^{ex} : n \in \text{RA} \cup \text{CA}\}|}{\textit{wordcount}}$$

- **Locution count** Number of L-nodes in the graph.

$$|\{n \in N^d : n \in \text{L}\}|$$

- **Support locution density** RA-nodes in the graph relative to number of L-nodes.

$$\frac{|\{n_i \in N : n_i \in \text{RA}\}|}{|\{n_j \in N^d : n_j \in \text{L}\}|}$$

- **Conflict locution density** CA-nodes in the graph relative to number of L-nodes.

$$\frac{|\{n_i \in N^{ex} : n_i \in \text{CA}\}|}{|\{n_j \in N^d : n_j \in \text{L}\}|}$$

- **Argument locution density** Combined RA and CA count relative to number of L-nodes.

$$\frac{|\{n_i \in N^{ex} : n_i \in \text{RA} \cup \text{CA}\}|}{|\{n_j \in N^d : n_j \in \text{L}\}|}$$

Structure-based analytics

The following analytics report on aspects of argument structure.

- **Serial argument count** Count of RA-nodes which participate in formation of a serial argument, i.e. count of RA-nodes whose premise is the conclusion of another RA, or whose conclusion is the premise of another RA.

$$\begin{aligned}|\{n_i \in N :\ & n_i \in \text{RA},\\ & \exists n_i, n_k : \langle n_j, n_i\rangle, \langle n_i, n_k\rangle \in E,\\ & \exists n_m : n_m \in \text{RA}, n_m \neq n_i,\\ & \text{at least one of } \langle n_m, n_j\rangle, \langle n_k, n_m\rangle \in E\}|\end{aligned}$$

- **Linked argument count** Count of RA-nodes which form a linked argument, i.e. count of RA-nodes with multiple incoming edges from I-nodes.

$$\begin{aligned}|\{n_i \in N :\ & n_i \in \text{RA},\\ & |\{n_j : n_j \in \text{I}, \langle n_j, n_i\rangle \in E\}| > 1\}|\end{aligned}$$

- **Convergent argument count** Count of RA-nodes which participate in formation of a convergent argument, i.e. count of RA-nodes which have the same conclusion as some other RA-node.

$$\begin{aligned}|\{n_i \in N : n_i \in \text{RA},\ & \\ & \exists n_j, n_k : n_j \in \text{RA}, n_k \in \text{I}, n_j \neq n_i,\\ & \langle n_i, n_k\rangle, \langle n_j, n_k\rangle \in E\}|\end{aligned}$$

- **Divergent argument count** Count of RA-nodes which participate in formation of a divergent argument, i.e. count of RA-nodes which have the same premise as some other RA-node.

$$\begin{aligned}|\{n_i \in N : n_i \in \text{RA},\ & \\ & \exists n_j, n_k : n_j \in \text{I}, n_k \in \text{RA}, n_k \neq n_i,\\ & \langle n_j, n_i\rangle, \langle n_j, n_k\rangle \in E\}|\end{aligned}$$

- **Serial argument percentage** Percentage of RA-nodes which participate in formation of a serial argument.

$$\frac{\begin{array}{c}|\{n_i \in N : n_i \in \text{RA}, \exists n_i, n_k : \langle n_j, n_i\rangle, \langle n_i, n_k\rangle \in E,\\ \exists n_m : n_m \in \text{RA}, n_m \neq n_i,\\ \text{at least one of } \langle n_m, n_j\rangle, \langle n_k, n_m\rangle \in E\}|\end{array}}{|\{n_l \in N : n_l \in \text{RA}\}|}$$

- **Linked argument percentage** Percentage of RA-nodes which form a linked argument.

$$\frac{|\{n_i \in N : n_i \in \text{RA}, |\{n_j : n_j \in \text{I}, \langle n_j, n_i\rangle \in E\}| > 1\}|}{|\{n_k \in N : n_k \in \text{RA}\}|}$$

- **Convergent argument percentage** Percentage of RA-nodes which participate in a convergent argument.

$$\frac{\begin{array}{l}|\{n_i \in N : n_i \in \text{RA},\\ \quad \exists n_j, n_k : n_j \in \text{RA}, n_k \in \text{I}, n_j \neq n_i,\\ \qquad \langle n_i, n_k\rangle, \langle n_j, n_k\rangle \in E\}|\end{array}}{|\{n_l \in N : n_l \in \text{RA}\}|}$$

- **Divergent argument percentage** Percentage of RA-nodes which participate in a divergent argument.

$$\frac{\begin{array}{l}|\{n_i \in N : n_i \in \text{RA},\\ \quad \exists n_j, n_k : n_j \in \text{I}, n_k \in \text{RA}, n_k \neq n_i,\\ \qquad \langle n_j, n_i\rangle, \langle n_j, n_k\rangle \in E\}|\end{array}}{|\{n_l \in N : n_l \in \text{RA}\}|}$$

- **Conflict source attacked** Count of I-nodes which are the attacker in a conflict (I-nodes with an outgoing edge to a CA) which are themselves attacked (additionally with an incoming edge from a CA).

$$\begin{aligned}|\{n_i \in N : n_i \in \text{I}\ & \\ & \exists \langle n_j, n_i\rangle, \langle n_i, n_k\rangle \in E,\\ & n_j, n_k \in \text{CA} \cap N^{ex}\}|\end{aligned}$$

- **Conflict source supported** Count of I-nodes which are the attacker in a conflict (I-nodes with an outgoing edge to a CA) which are themselves supported (additionally with an incoming edge from an RA).

$$\begin{aligned}|\{n_i \in N : n_i \in \text{I}\ & \\ & \exists \langle n_j, n_i\rangle, \langle n_i, n_k\rangle \in E,\\ & n_j \in \text{RA}, n_k \in \text{CA} \cap N^{ex}\}|\end{aligned}$$

- **Premise count** Count of I-nodes which serve as premise to some argument, i.e. count of I-nodes with an outgoing edge to an RA-node.

$$|\{n_i \in N : n_i \in \text{I}, \exists n_j \in \text{RA}, \langle n_i, n_j\rangle \in E\}|$$

- **Conclusion count** Count of I-nodes which serve as conclusion to some argument, i.e. count of I-nodes with an incoming edge from an RA-node.

$$|\{n_i \in N : n_i \in \text{I}, \exists n_j \in \text{RA}, \langle n_j, n_i\rangle \in E\}|$$

- **Support-Attack ratio** The number of support relations divided by the number of attack relations.

$$\frac{|\{n_i \in N : n_i \in \mathrm{RA}\}|}{|\{n_j \in N^{ex} : n_j \in \mathrm{CA}\}|}$$

- **Premise-Conclusion ratio** The number of nodes which act as a premise for some support relation divided by the number of nodes which act as a conclusion for some support relation.

$$\frac{|\{n_i \in N : n_i \in \mathrm{I}, \exists n_j \in \mathrm{RA}, \langle n_i, n_j \rangle \in E\}|}{|\{n_i \in N : n_i \in \mathrm{I}, \exists n_j \in \mathrm{RA}, \langle n_j, n_i \rangle \in E\}|}$$

Language-based analytics

Let $\mathtt{past}(n_i)$, $\mathtt{present}(n_i)$ and $\mathtt{base}(n_i)$ be the past, present and base verb-form scores of t_{n_i}, the text of node n_i.

- **Past tense score** Average past tense score of I-nodes in the graph.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{past}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

- **Present tense score** Average present tense score of I-nodes in the graph.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{present}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

- **Base verb score** Average base verb score of I-nodes in the graph.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{base}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

Let $\mathtt{pos}(n_i)$, $\mathtt{neu}(n_i)$, $\mathtt{neg}(n_i)$ and $\mathtt{comp}(n_i)$ respectively be the positive, neutral, negative and compound sentiment scores of t_{n_i}, the text of node n_i.

- **Average argument component positive sentiment score** Average positive sentiment score of I-nodes in the argument structure.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{pos}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

- **Average argument component neutral sentiment score** Average neutral sentiment score of I-nodes in the argument structure.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{neu}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

- **Average argument component negative sentiment score** Average negative sentiment score of I-nodes in the argument structure.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{neg}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

- **Average argument component compound sentiment score** Average compound sentiment score of I-nodes in the argument structure.

$$\frac{\forall n_i \in N : n_i \in \mathrm{I}, \sum_{i=1}^{n} \mathtt{comp}(n_i)}{|\{n_i \in N : n_i \in \mathrm{I}\}|}$$

Let *concat*($\{n_i, \ldots, n_j\}$) be some concatenation of $t_{n_i}, \ldots, t_{n_j}$, the texts of nodes $n_i, \ldots, n_j$. Note that for the purposes of sentiment scoring, the order in which the texts are concatenated to form a single text is arbitrary.

- **Argument structure positive sentiment score** Positive sentiment score for the concatenated text of all I-nodes in the argument structure.

$$\mathtt{pos}(\textit{concat}(\{n_i \in N : n_i \in \mathrm{I}\}))$$

- **Argument structure neutral sentiment score** Neutral sentiment score for the concatenated text of all I-nodes in the argument structure.

$$\mathtt{neu}(\textit{concat}(\{n_i \in N : n_i \in \mathrm{I}\}))$$

- **Argument structure negative sentiment score** Negative sentiment score for the concatenated text of all I-nodes in the argument structure.

$$\mathtt{neg}(\textit{concat}(\{n_i \in N : n_i \in \mathrm{I}\}))$$

- **Argument structure compound sentiment score** Compound sentiment score for the concatenated text of all I-nodes in the argument structure.

$$\mathtt{comp}(\textit{concat}(\{n_i \in N : n_i \in \mathrm{I}\}))$$

Hypothesis-comparison analytics

Two classes of analytic are used to capture differences among the hypotheses within a forecast: standard deviation and ratio.

Let the set of hypotheses in forecast F be $H = \{h_1, \ldots, h_n\}$, some hypothesis be $h_j \in H$, and some analytic be λ_i, where λ_{i,h_j} is the value of analytic λ_i for hypothesis h_j . The standard deviation analytic σ_{λ_i} for forecast F and analytic λ_i is then the standard deviation of the values of that analytic for each hypothesis within the forecast.

$$\sigma_{\lambda_i} = \sigma\{\lambda_{i,h_1}, \ldots, \lambda_{i,h_n}\}$$

A ratio analytic for a forecast consists of the ratio between the smallest value for that analytic among the hypotheses and the largest value for that analytic among the hypotheses. Where Λ_i is the set of values of λ_i for all

$$\mathrm{ratio}_{\lambda_i} = \frac{\lambda_{h_i} : h_i \in H, \nexists h_j : h_j \in H, \lambda_{h_j} < \lambda_{h_i}}{\lambda_{h_k} : h_k \in H, \nexists h_l : h_l \in H, \lambda_{h_l} > \lambda_{h_k}}$$

The analytics described above for which hypothesis standard deviation and ratio analytics are calculated for forecasts are as follows:

- **Support count**
- **Conflict count**
- **Linked argument count**
- **Linked argument percentage**
- **Convergent argument count**
- **Convergent argument percentage**
- **Divergent argument count**
- **Divergent argument percentage**
- **Serial argument count**
- **Serial argument percentage**
- **Conflict source attacked**
- **Conflict source supported**
- **Premise count**
- **Conclusion count**
- **Support-Attack ratio**
- **Premise-Conclusion ratio**

Four analytics are additionally provided which use an adjusted notion of wordcount. In order to adapt the notion of word-density to hypotheses, a wordcount is used based on the total wordcount across information nodes in the graph, rather than original text. Where wordcount(n_i) indicates the wordcount of t_{n_i}, the text of node n_i, these additional/adjusted analytics for which standard deviation and ratio analytics are calculated are:

- **Wordcount** The summed wordcount of all I-nodes in the graph, *wordcount*$_{\mathrm{I}}$.

$$\text{For } \{n_i \in N : n_i \in \mathrm{I}\}, \sum_{i=1}^{n} \text{wordcount}(n_i)$$

- **Support word density** RA count relative to *wordcount*$_{\mathrm{I}}$ calculated from I-nodes of the graph.

$$\frac{|\{n \in N : n \in \mathrm{RA}\}|}{\textit{wordcount}_{\mathrm{I}}}$$

- **Conflict word density** CA count relative to *wordcount*$_{\mathrm{I}}$ calculated from I-nodes of the graph.

$$\frac{|\{n \in N^{ex} : n \in \mathrm{CA}\}|}{\textit{wordcount}_{\mathrm{I}}}$$

- **Argument word density** Combined RA and CA count relative to *wordcount*$_{\mathrm{I}}$ calculated from I-nodes of the graph.

$$\frac{|\{n \in N^{ex} : n \in \mathrm{RA} \cup \mathrm{CA}\}|}{\textit{wordcount}_{\mathrm{I}}}$$

Lastly, five analytics were produced which were only used in the comparison analytics. The underlying analytic definitions for which the additional standard deviation and ratio analytics are calculated are as follows:

- **Restating count** Count of uses of *Restating* in the graph.

$$|\{n_i \in N^d : t_{n_i} = \text{`Restating'}\}|$$

- **Undercut count** Count of CA-nodes which indicate an undercut-type attack, i.e. count of CA-nodes with an outgoing edge to an RA-node.

$$|\{n_i \in N^{ex} : n_i \in \mathrm{CA}, \exists \langle n_i, n_j \rangle \in E : n_j \in \mathrm{RA}\}|$$

- **Rebuttal count** Count of CA-nodes which indicate a rebuttal-type attack, i.e. count of CA-nodes with an outgoing edge to an I-node.

$$|\{n_i \in N^{ex} : n_i \in \mathrm{CA}, \exists \langle n_i, n_j \rangle \in E : n_j \in \mathrm{I}\}|$$

- **Max support chain** The longest path via RA- and I-nodes in the argument graph in a single direction. That is, the length of the argument with the structurally longest sequence of inferences. Let $P_{a_{RA}}$ be the set of paths in argument a such that: $\forall \langle n_i, n_j \rangle \in P_{a_{RA}} : n_i, n_j \in \mathrm{I} \cup \mathrm{RA}$.

$$|p_i| : p_i \in P_{a_{RA}}, \nexists p_j \in P_{a_{RA}} : |p_i| < |p_j|$$

- **Max conflict chain** The longest path via CA- and I-nodes in the argument graph in a single direction. That is, the length of the conflict sequence with the structurally longest sequence of conflicts. Let $P_{a_{CA^{ex}}}$ be the set of paths in argument a such that: $\forall \langle n_i, n_j \rangle \in P_{a_{CA^{ex}}} : n_i, n_j \in \mathrm{I} \cup \mathrm{CA}, n_i, n_j \in N^{ex}$.

$$|p_i| : p_i \in P_{a_{CA^{ex}}}, \nexists p_j \in P_{a_{CA^{ex}}} : |p_i| < |p_j|$$