Annotation

Dear annotators,

Arguments may be written in a wide variety of forms and may not always be directly apparent. Likewise, an **Argument** may address one topic or more, be rhetorically and linguistically heavy, or the inferences stated and reasons given may not be immediately obvious. Nevertheless, the individual text passages in a tweet can form an **Argument** in their effect. It is not important for you whether the **Argument** is right or wrong or whether it corresponds to your personal world view and opinion. To simplify this task, we differentiate between tweets that contain an inference as a key component of an **Argument** and those that do not.

To identify **Argument** constituents, we use the Cambridge Dictionary:

- **Inference**: *a guess that you make or an opinion that you form based on the information that you have*.
- **Information**: *facts or details about a person, company, product, etc.*.

Tweets that make an **Argument** can be classified as either:

- **Statement**: a tweet where only **Inference** is presented like *something that someone says or writes officially, or an action done to express an opinion*.
- **Reason**: a tweet where the **Inference** is based on **Information** mentioned in the tweet such as a source-reference or quotation, and thus reveals the author's motivation *to try to understand and to make judgments based on practical facts*.

Tweets that make **No-Argument** can be classified as either:

- **Notification**: a tweet that limits itself to only provide information like media channels promoting their latest articles.
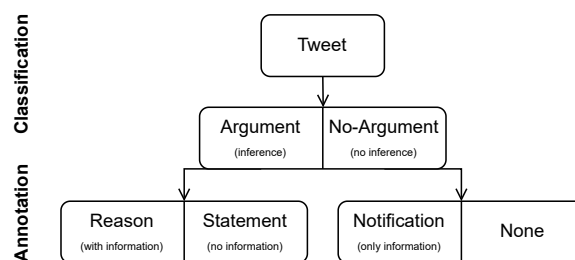- **None**: a tweet having neither inference or information, including hate-speech or spam.

Figure 1: Hierarchy of Argument Mining on Twitter.

Before answering the questions ask yourself:

1 What does the author of the tweet want you to believe (**Inference**)?
2 What does the author of the tweet want you to know (**Information**)?
3 Does (1) emerge from (2)?

In the case of uncertainty consider the following:

1 Look the tweet up on Twitter
2 Follow the conversation down to the tweet.
3 *Concrete dates* are **Information**
4 *Quotes* and *headlines* are **Information**
5 *Experiences* are **Information**
6 *Considerations* are **Information**
7 *Standpoints* are **Inference**
8 *Rhetorical questions* are **Inference**
9 *Hashtags can* provide **Inference**
10 *Insults are no* **Inference**
11 *Exceptions might indicate* **Inference**
12 Tend to **0** in case of absolute doubt

Task:

1 Detect **Information** and **Inference**.
Mark each identified component with **1** else use **0**

2 Specify if you have strong concerns about interpreting the tweet too strongly and selecting the component too artificially? Therefore, how well did the components emerge? Mark the **Difficulty** with:

1: Easy (The component(s) is/are directly visible)
2: Normal (Repeated consideration reveals the component(s))
3: Hard (Strong concerns about any component(s) present)

Examples:

Tweet 1 (Reason):

> The formula:
> Not everyone who voted Leave is racist.
> But everyone who's racist voted Leave.
> Not everyone who voted Leave is thick.
> But everyone who's thick voted Leave.
> The thick racists therefore called the shots,
> whatever the thoughts of the minority of others.
> #thick #Brexit

>   **Topics**: Brexit, Racism, Minority, Leave, Thick
>   **Inference**: 1
>   **Information**: 1
>   **Difficulty**: 1

In this example, **Information** and **Inference** are chosen because, on the one hand, a consideration is made, and on the other hand, a concrete image of Brexit is drawn. Specifically, the consideration *"everyone who's racist voted Leave"* and *"everyone who's thick voted Leave"* leads to the **Inference** (*"The thick racists [...] called the shots"*) relative to the exceptions *"Not everyone who voted Leave is racist"* and *"Not everyone who voted Leave is thick"*. Likewise, it alludes to the lack of hearing minorities, which is intended to emphasize the perceived backward step in Brexit (#*thick #Brexit*). Obviously, this tweet hints at the right-winged aspects of Brexit and reminds of the negative (*"thick"*) characteristics of marginalization as they are often present in racism and right-winged movements. For the entirety of the tweet, it is not relevant here whether the **Inference** and **Information** are true or false but contribute to the **Argument**.

Tweet 2 (Notification):

> #Mexico top court declares criminalizing #Abortion_06 unconstitutional-JURIST-News Mexico's Supreme Court of Justice of the Nation ruled Tuesday that total #criminalization of #abortion is unconstitutional.
> #AbortoLegalMexico
> #USSupremeCourt
> #SupremeCourt
> `https://t.co/xLj5PZijOL`

>   **Topics**: Mexico, Abortion, Supreme Court, USA
>   **Inference**: 0
>   **Information**: 1
>   **Difficulty**: 1

In this example, only a summary of the article in the link below is presented in the tweet. To see this, the URL to the article must be followed. Likewise, the tweet summarizes and recites the main message of the article, as it also occurs in the headline.

Tweet 3 (Statement):

Possible tangible benefit of #Brexit: a fairer immigration system.

> **Topics**: Immigration, Brexit, Benefit, Possibility
> **Inference**: 1
> **Information**: 0
> **Difficulty**: 1

In this Tweet, only an **Inference** in the form of an assertion is made from anticipated and implicit **Information** (*the immigration system is unfair*). Any further elements are missing; the assertion is assigned to the topic in a substantial but challenging role.

Tweet 4 (None):

@sinnfeinireland Blah blah blah blah blah blah

> **Topics**: Ireland → Brexit (?)
> **Inference**: 0
> **Information**: 0
> **Difficulty**: 1

After a brief consideration, the individual components for an **Argument** are to be strongly doubted.

Tweet 5 (Reason):

Love my new favorite abortion shirt! Abortions are as good as, if not better than, all medical procedures! Get yourself this badass shirt and donate to @TEAFund: `https://t.co/vSiwJaMZjh` #abortion #AbortionBan #AbortionIsHealthcare #txlege #texas `https://t.co/5iTXpODVIv`

> **Topics**: Abortion, Texas, Healthcare, Donation, Abortion-Shirt
> **Inference**: 1
> **Information**: 1
> **Difficulty**: 2

This tweet primarily looks like an advertisement and call for donations. However, on closer reading, both main components are present as well. Here, the **Information** is provided by following the first URL to an information page or by the consideration *"Abortions are as good as, if not better than, all medical procedures"*. Where the consideration appears to have priority over the information page. Likewise, the **Inference** is in the hashtag *#AbortionIsHealthcare*. In both cases, the components of an **Argument** become apparent after short consideration (**Difficulty: 2**).