**TODO:**

**# Add column of how previous work maps onto mine**

**# Send Kevin rationale for annotation**

**Scientific goals: Evidence and New Topic**

**Task:** Please annotate each sentence as evidence, new topic, or neither as defined below. Please also include the trigger word(s) that led you to the conclusion that the sentence was evidence or new topic, and the scientific goal that comes from that sentence specifically.

<SML: maybe include an example here, or after your 1,2,3 list, especially what you mean by ‘scientific goal that comes’>

General caveat: Some lexical cues do not connote the type of annotation.

1. Evidence annotation = **needs more evidence**
   1. Sentences where the author is not definitive on the truth value of the statement, including hedges. The truth value is not determined because more investigation is necessary. Ask yourself if more investigation is needed or not.
      1. Potential lexical cues: suggests, may, might, should, suspect, suppose, seem, appear, likely, possible, probably, what if, speculate, propose, what if, etc.
      2. Other lexical cues:
   2. We assume a statement annotated as “evidence” needs **more** **evidence** in order to fully determine the truth value of the statement.
   3. **Goal for knowledge: acquire evidence**
2. New topic annotation = **unexplored topic**
   1. Sentences that introduce a topic for future investigation. There is little known information about the topic presented except for the fact that it has not been studied.
      1. Potential lexical cues: unclear, little is known, unexplored, open question, not been determined, not been explored, not been studied, controversial, debated, further investigation, not assessed, not comprehensively, limitation, not designed to, did not allow for, etc.
   2. We assume a statement annotated as “new topic” could lead to another paper on the topic.
   3. **Goal for knowledge: topic in** **need of investigation**
3. Neither annotation
   1. Neither of the other 2

**Similar tasks**

|  |  |  |
| --- | --- | --- |
| **Author** | **Task** | **Definition** |
| Elizabeth White Thesis | Argumentation | Summarizing the paper = scientific methods, cognition, discourse, negation, causation, and modality |
| Zerva/Ananiado (2017) | Uncertainty | Assessing the confidence of related information in terms of the certainty of a statement based on its textual context – hypotheses, speculated outcome, case under investigation, result attributed to an unclear external source (Medlock 2008)   * Uncertainty of events and particularly linking it with interactions in pathways and interaction networks |
| Light (2004) - future work/applications:   * Speculation search engine * Knowledge discovery test set | Speculative language – distinguished between high and low but didn’t work | Expressions of levels of belief: the expressions of hypotheses, tentative conclusions, hedges, and speculations – (Affect) |
| Vincze and Csirik (2008)   * Annotation guidelines includes keywords but cautions * Scientific texts harder than clinical stuff – not many articles annotated | Bioscope – uncertainty and negation corpus | Impressions, hypothesized explanations of experimental results or negative findings  Speculation = hedge/soft negation |
| Kilicoglu and Bergler (2008) | Speculative language = hedging | Nice review of previous work  Builds off of hyland |
| Hyland (1998) | Hedging | One part of epistemic modality, it indicates an unwillingness to make an explicit and complete commitment to the truth of propositions |
| Medlock (2007)   * Single terms as features, based on intuition that many hedge cues are single terms * Defined non-hedges as no cues for hedges = problematic   Future = | Hedging | Under the umbrella of subjectivity   * Authorial opinion   Hedge:   * an assertion relating to a result that does not necessarily follow from work presented but extrapolated from it (Light) * relay of hedge made in previous work * statement of knowledge paucity * speculative question * statement of speculative hypothesis * anaphoric hedge reference |
| Farkas - CoNLL (2010)   * application: information extraction, making sure everything is certain | Uncertainty cues and their linguistic scope | Hedges = indicating that authors do not or cannot back up their opinions/statements with facts |
| Ganter (2009) | Hedges by chasing weasels! – Wikipedia   * goal: to get rid of non-factual information | Offer an opinion without really backing it up and… are really used to express a non-neutral point of view |
| Ram and Hunter (1992) | Goals/desires for knowledge | A knowledge goal represents the need to fill in gaps in the reasoner’s knowledge base that are detected when a piece of information required for a task turns out to be missing, incorrect, or otherwise problematic   * gaps give rise to new questions, which in turn stimulate further interest in the topic * underlying goal: learn and improve one’s model of the world |

**Ignorance identifying guidelines – SCIENTIFIC GOALS - old**

Determine if a sentence/passage given includes some ignorance. Ignorance is defined as a known unknown in the literature throughout all time that a goal can be created for to solve. Ignorance is a form of incompleteness that the goal aims to complete. For example, “consequently, an unknown proportion of the current plant diversity in this habitat type will go extinct if no new conservation actions aimed at large-scale habitat restoration are initiated.” The goal here is to initiate new conservation actions aimed at large-scale habitat restoration.

In contrast to an ignorance statement is a definite statement including observations, methods, and previous work usually (light2004language). If no scientific goal can come from the statement, then it is not ignorance

There is gradation in terms of (in)completeness: how complete is the question we are trying to answer. For example, “these results suggest that lipolytic enzymes could be regarded as potent targets for future drug development.” The goal is to find out if lipolytic enzymes can be used for drug development. The “suggests” shows that we have some knowledge to support this idea but need more before we conclude that we can use it.

**Task:** Mark each sentence with an I for ignorance, with a N for not ignorance, or a U for unclear. If the sentence is marked ignorance, please highlight the word(s) that indicated that and write out the goal that needs to be solved.