Babble Labble: Learning from Natural Language Explanations

Braden Hancock, Paroma Varma, Percy Liang, Chris Ré

Big Idea: Explain Yourself!

- Collecting supervision one bit (one binary label) at a time is tedious, slow, expensive, and not very scalable.
- Explanations describing why certain labels were given can be used to vicariously label large amounts of unlabeled data.
- The potentially overlapping/conflicting labels can be resolved statistically to generate probabilistic "noise-aware" labels for downstream discriminative models.

Explain why how you labeled this…

"Bob and his wife Alice visited Stanford University last Thursday."

Do you think Bob and Alice are married?

- (●) Yes
- ( ) No

Why?

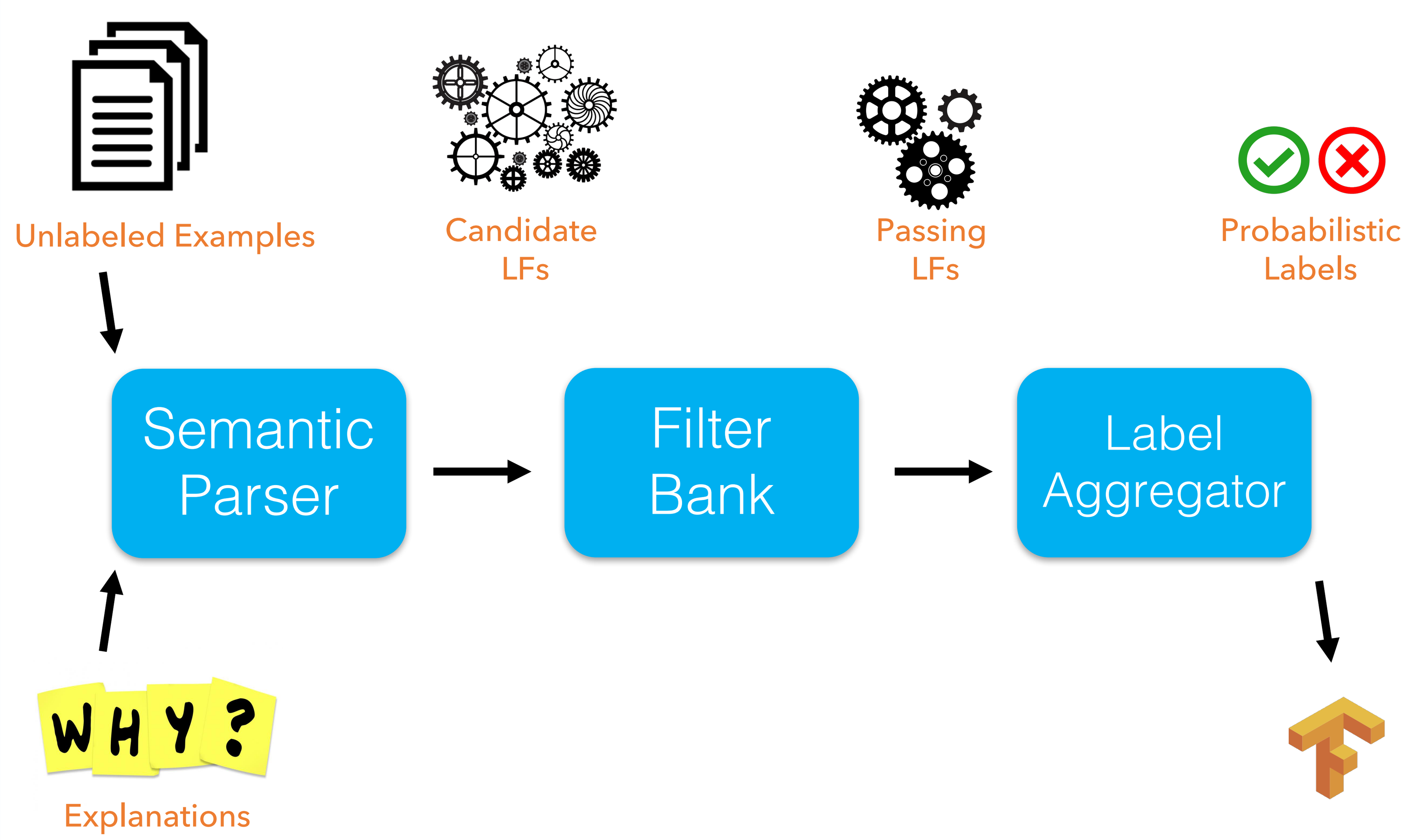
Because the words 'his wife' are immediately to the left of Alice.

And get supervision on these for free…

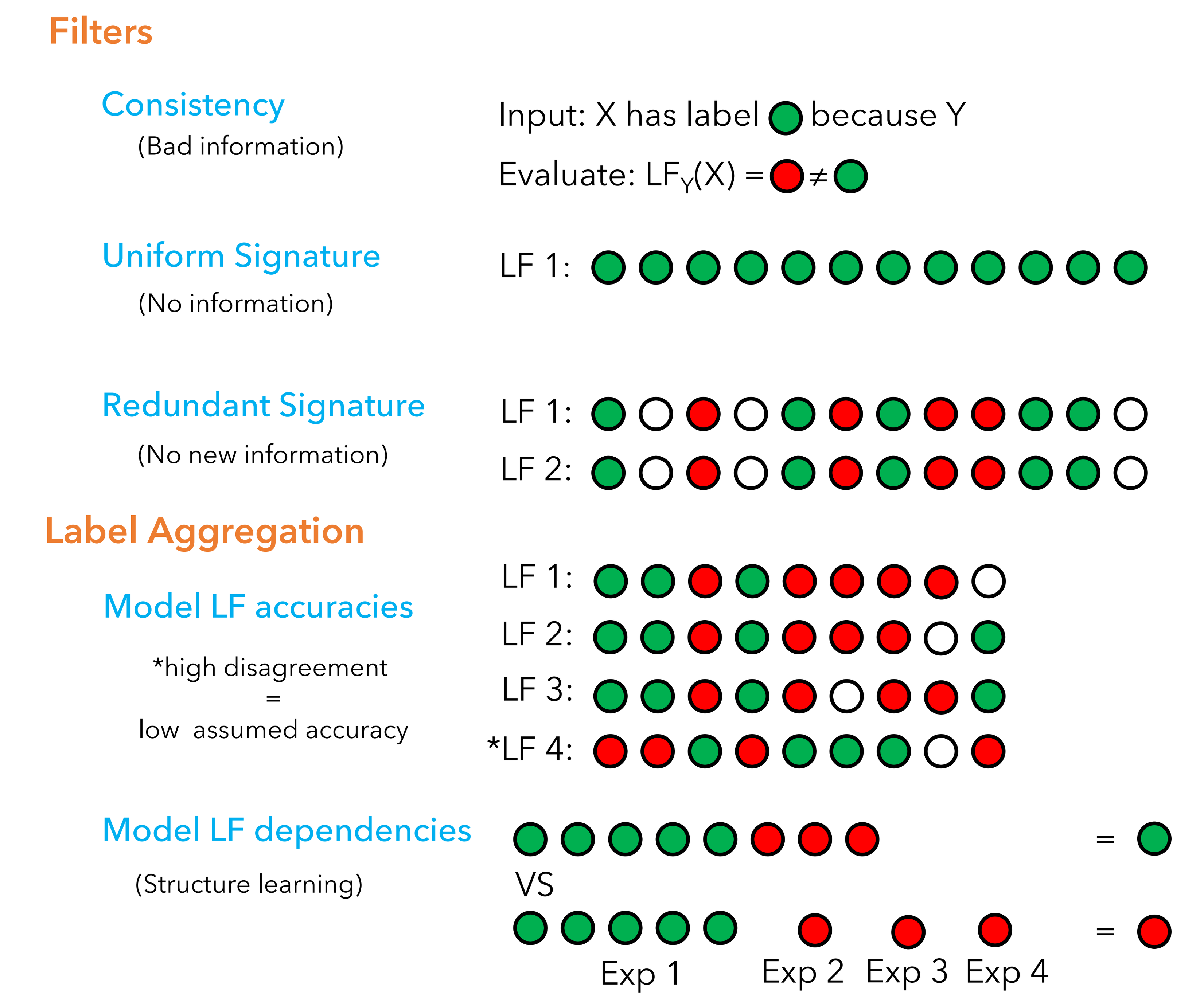
- ✔ "Barack batted back tears as he thanked his wife, Michelle, for all her help."
- ✔ "Both Bill and his wife Hillary smiled and waved at reporters as they rode by."
- ✔ "George attended the event with his wife, Laura, and their two daughters."

The Babble Labble Framework

Key Idea: Use natural language explanations to programmatically label large training sets. Use filters and noise modeling to address the ambiguity introduced by using natural language.



Filter Bank & Label Aggregator



Motivation

Why supervise with natural language instead of formal programs?

1) Ease of Use:

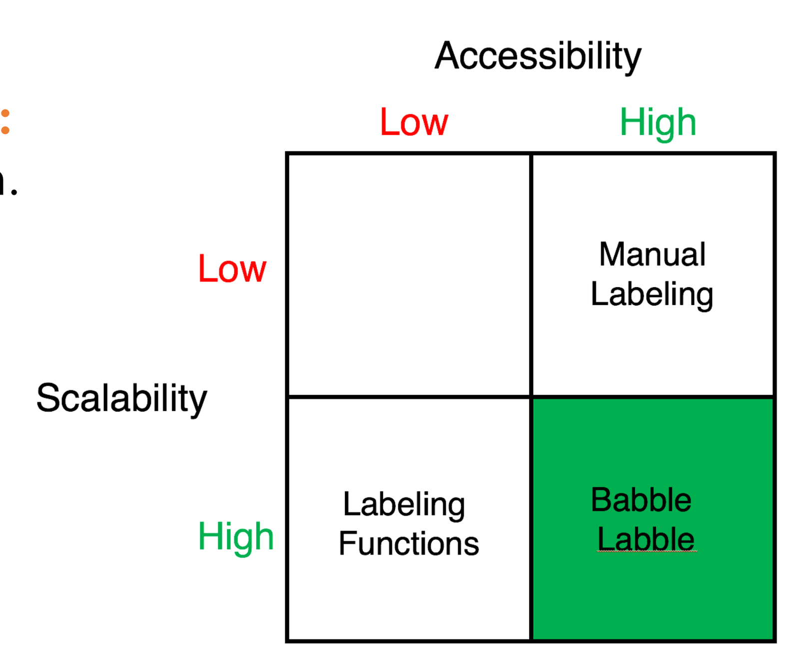
Conversational interfaces are the new normal!

2) Faster supervision:

No need to look up syntax or find missing semicolons!

3) More sources of supervision:

Not everyone can program.
Everyone can explain.



Semantic Parser

Example X:

Barack (A) and his wife Michelle (B) live in the White House.

Explanation Y:

"'wife' precedes Michelle in the sentence"

Normalized Explanation:

"'wife' precedes B in the sentence"

Candidate Programs:

#	Program	Observations
1	Before('wife', B)	Correct
2	Before(B, 'wife')	False for X
3	In(B, Sentence)	Trivially True for all examples
4	After(B, 'wife')	Correct, but redundant with #1

Traditional approach: use labeled examples to learn to rank candidate programs
Our approach: handle parsing ambiguity downstream with filters and modeling

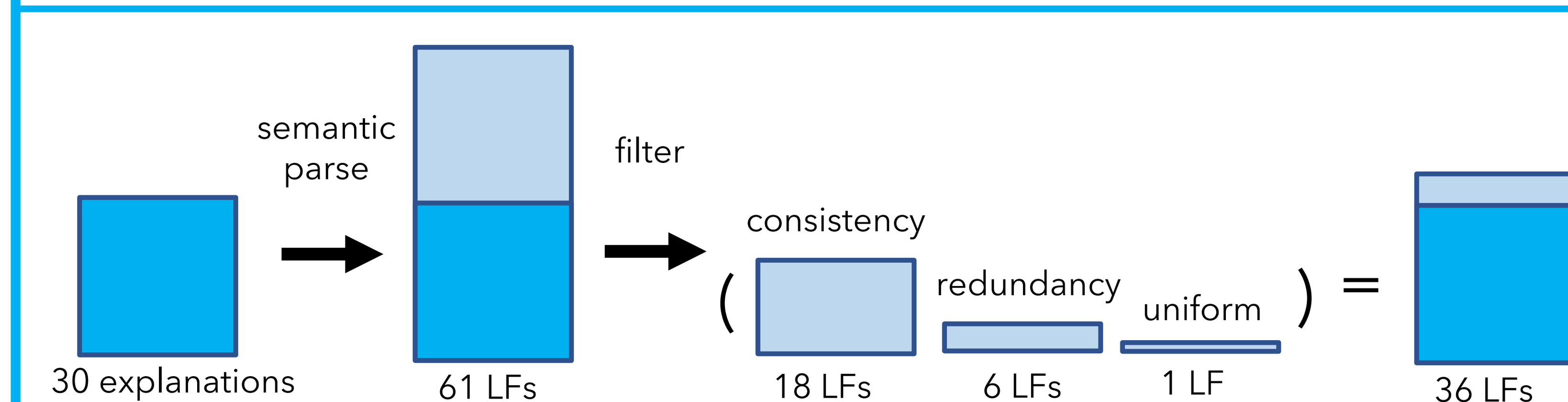
Experimental Results: Chemical-Disease Relations

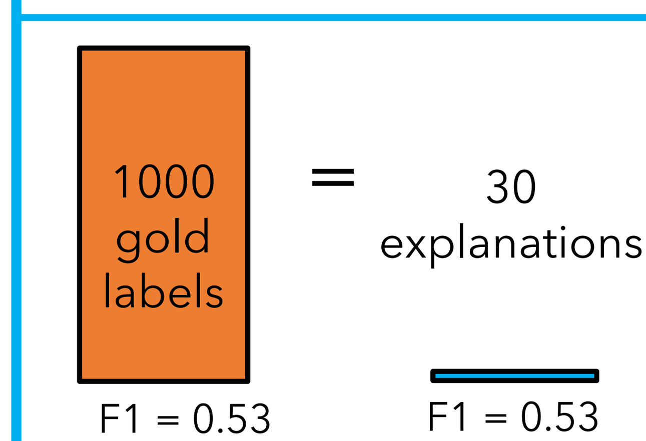
Task: Identify mentions of chemicals causing diseases in PubMed.

True – "because 'induced by', 'caused by', or 'due to' is between the disease and the chemical"

False – "because a treatment word is between the chemical and the disease and the chemical is within 100 characters to the left of the disease"

False – "because the pair of canonical IDs of the chemical and the disease is in the therapeutic combinations dictionary"





Same quality with 33x fewer inputs!