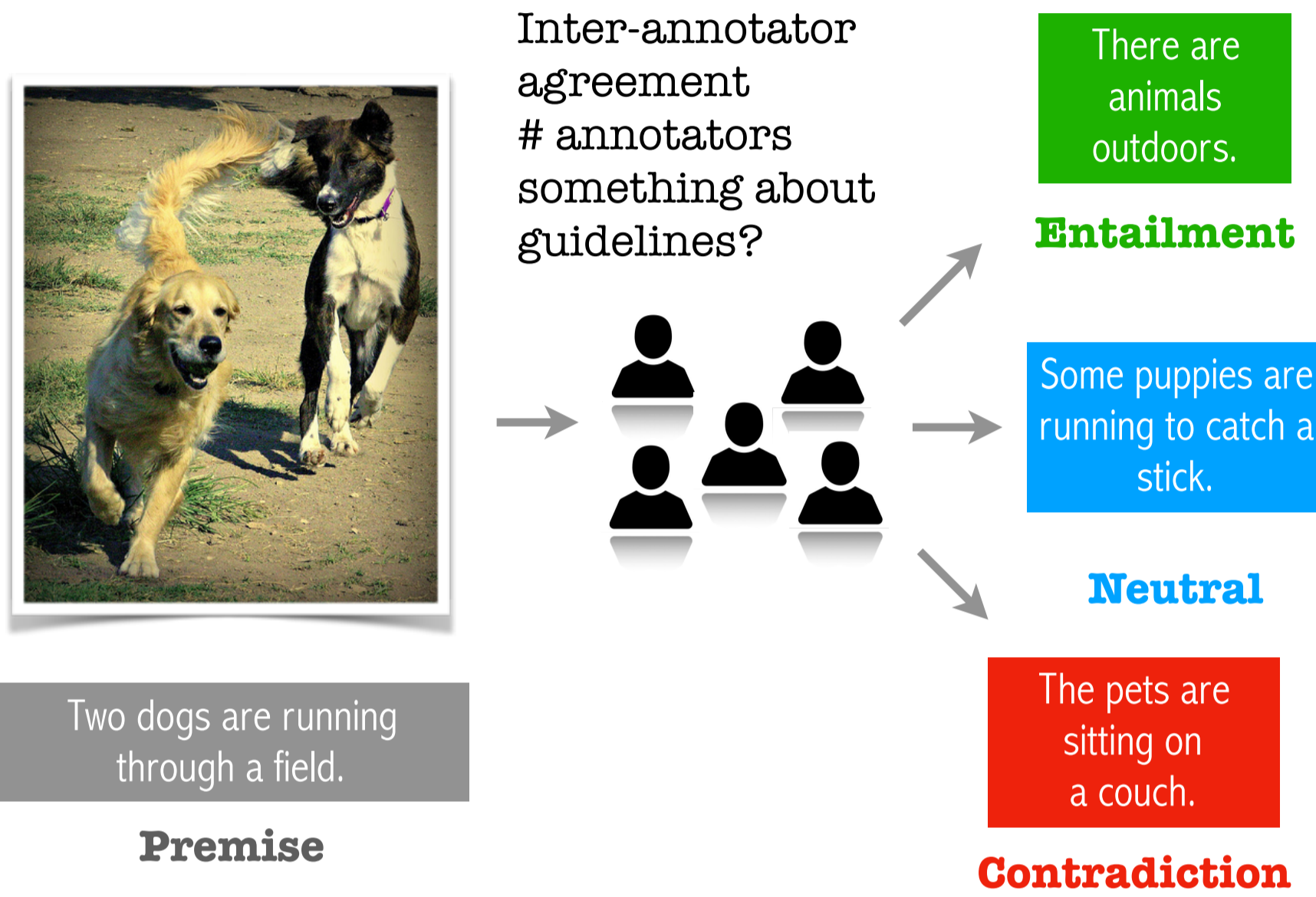
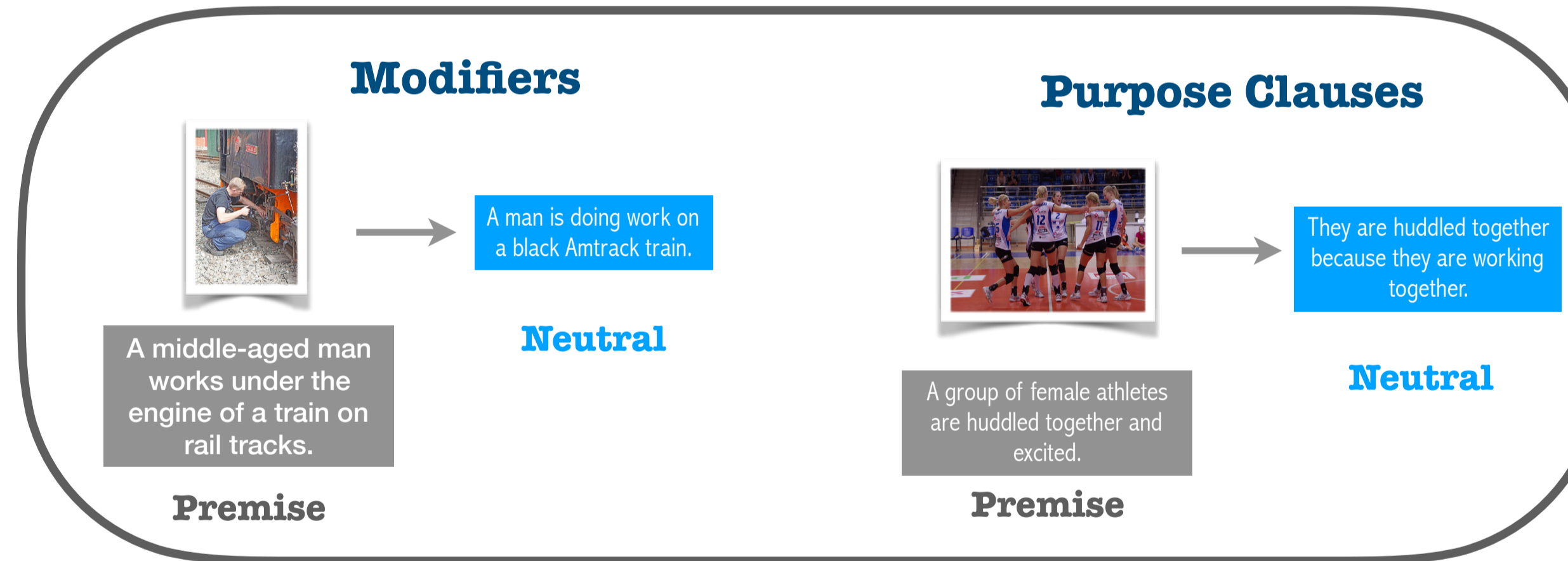
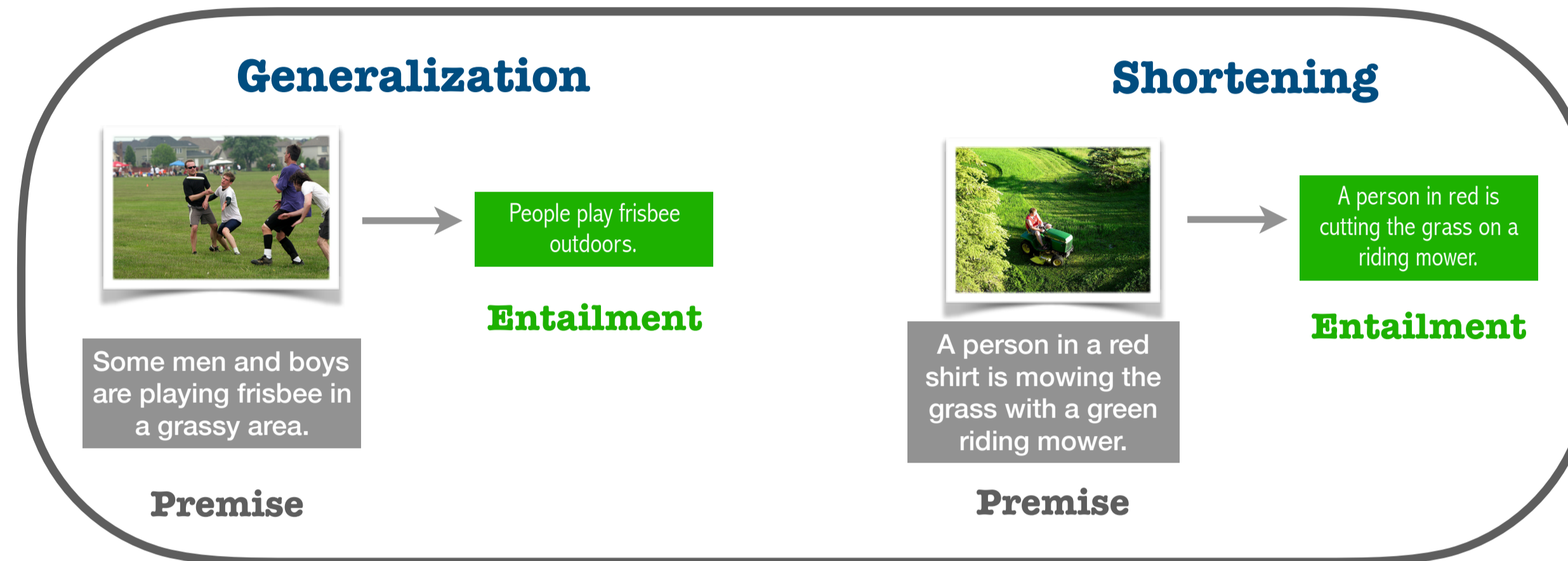


## Natural Language Inference

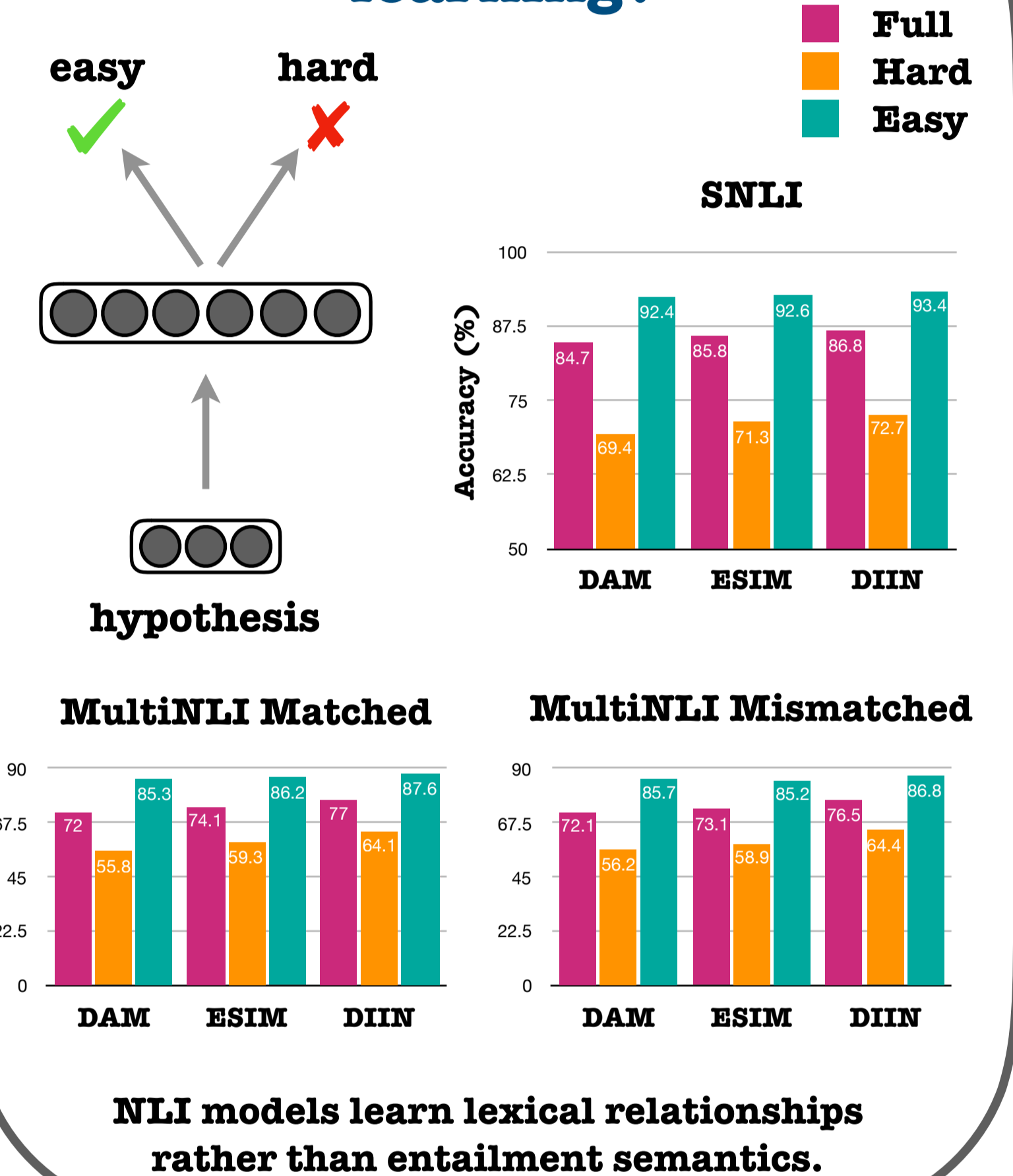


SNLI (Bowman et. al., 2015)  
MultiNLI (Williams et. al., 2017)

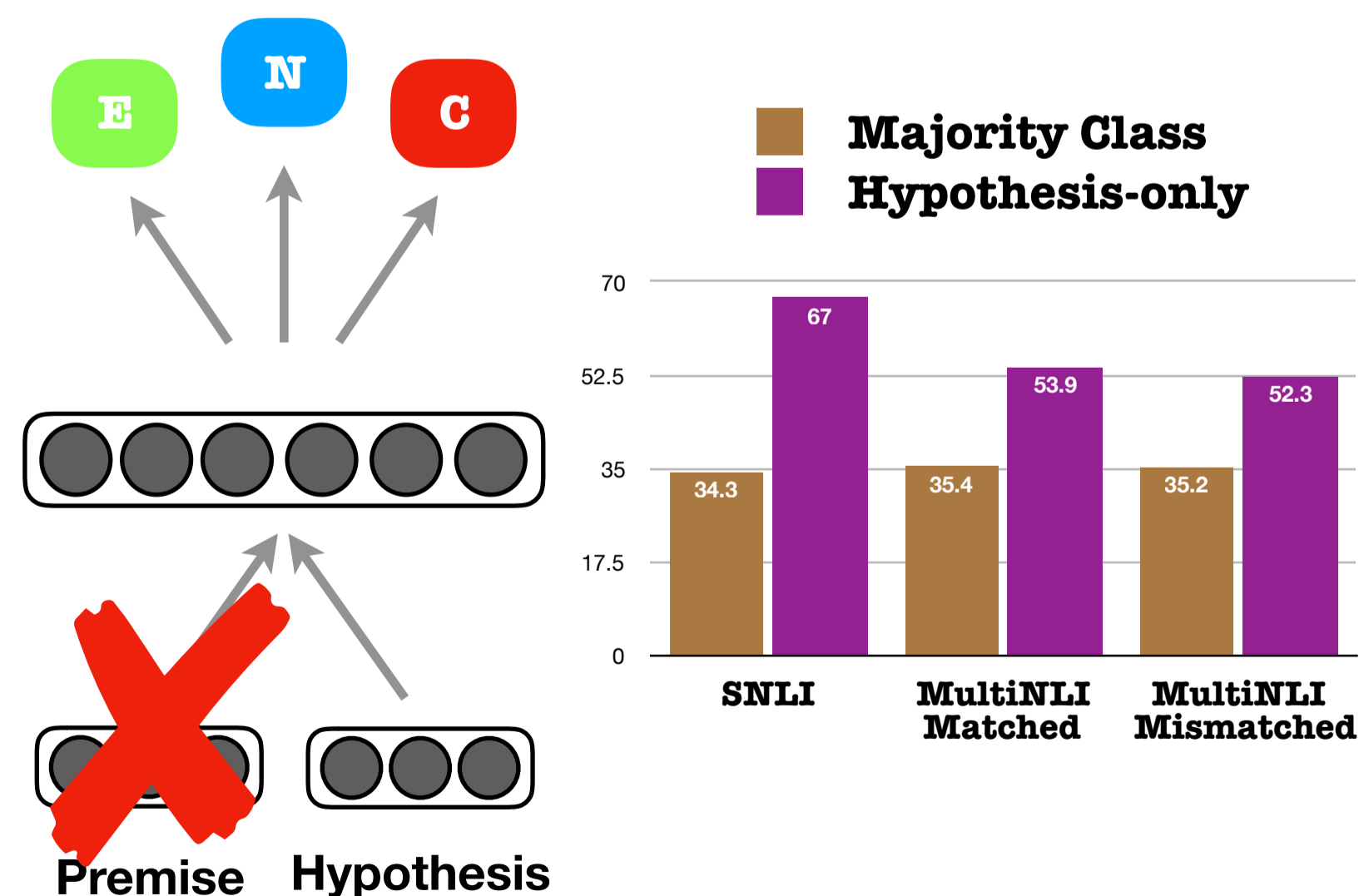
## Annotation Artifacts



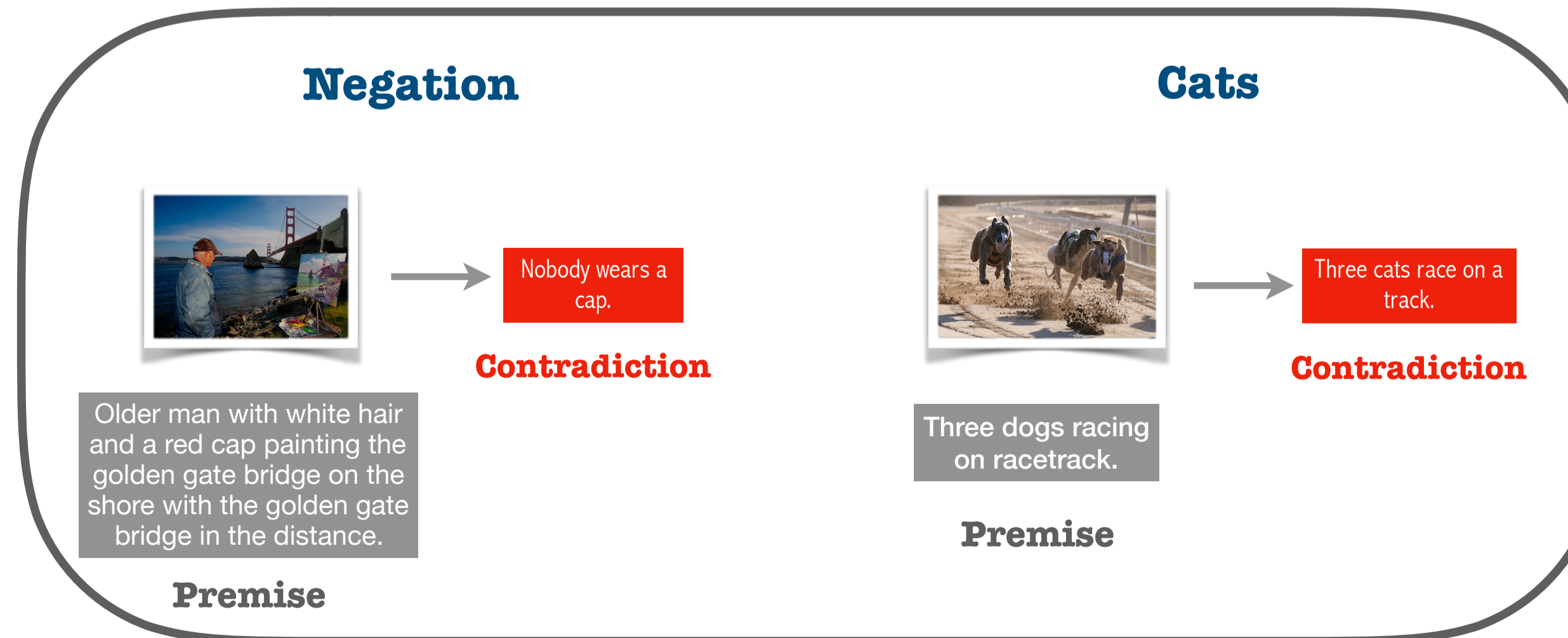
## What are NLI models really learning?



## Kicking out Premises...



fastText (Joulin et. al., 2016)



## Takeaways

1. These results are consistent with numerous findings of issues with NLP datasets.
2. Annotation artifacts can be addressed by improving annotation protocols to correct for these biases.
3. We release the Hard benchmarks on Kaggle and the SNLI website.