Understanding Negation in Positive Terms

Using Syntactic Dependencies (Published in EMNLP 2016)

Zahra Sarabi & Eduardo Blanco

University of North Texas

Contact Information:
Computer Science Department
University of North Texas
Denton, Texas

Email: zahrasarabi@my.unt.edu
eduardo.blanco@unt.edu
Website: zahrasarabi.com

Motivation

Negation often conveys positive meaning.

- **Most jobs now** don't last for **decades**.
  - **Few jobs** now last for decades.
  - Most jobs **in the past** lasted for decades.
  - Most jobs now last for **a few years.**

In this work we present a methodology to extract positive interpretations from a negative sentence, as intuitively done by human.

Main Objectives

1. Create a corpus of negation and their positive interpretations
   (a) Automatic generation of potential positive interpretations
   (b) Manual validation
2. Learning to score potential positive interpretations

Corpus Creation

- Two steps:
  1. Generate potential positive interpretations automatically using syntactic dependencies
  2. Validate potential positive interpretations (manual annotations)

Step 1. Generating Potential Positive Interpretations

Selecting negation

- select 8,168 **verbal negations** from OntoNotes
- verbal negation: tokens whose syntactic head is a verb and dependency type ***neg***

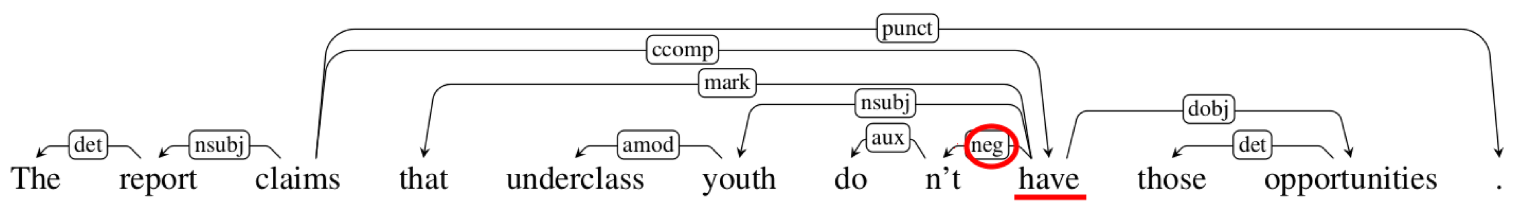

Figure 1: Verbal negation

Converting negations into their positive counterparts

1. Remove the negation mark
2. Remove auxiliaries, expand contractions, and rewrite third-person singular and past tense
3. Rewrite negatively-oriented polarity-sensitive items

Selecting relevant tokens

- Simplify the original statement by including only the negated verb and all tokens reachable from the negated verb traversing dependencies.

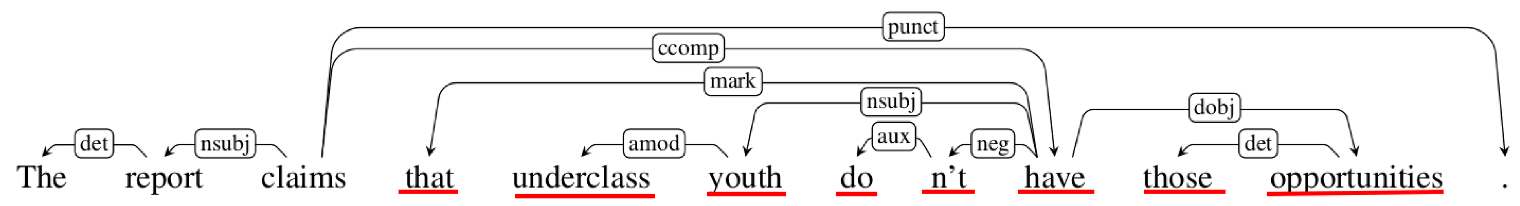

Figure 2: Selecting relevant tokens

after step 1.2	The report claims that underclass youth have those opportunities.
after step 1.3	Underclass youth have those opportunities.

Table 1: Exemplify steps 1.2 and 1.3

Generate potential interpretations from positive counterpart

1. Traverse the dependency tree from the negated verb and select all subtrees up to depth 3
2. Discard useless potential foci, e.g., whose syntactic dependency is *aux*
3. Rewrite each focus with *"someone/something/etc."*, appending *"but not text_of_focus"* at the end

Dependency		Underclass youth have those opportunities.
nsubj	coarse	**[Some people]** have those opportunities, but not ***Underclass youth***.
amod	fine	**[Some adjective]** youth have those opportunities, but not ***Underclass*** *youth*.
nsubj	fine	Underclass **[people]** have those opportunities, but not *Underclass* ***youth***.
dobj	coarse	Underclass youth have **[something]**, but not ***those opportunities***.
det	fine	Underclass youth have **[some]** opportunities, but not ***those*** *opportunities*.
dobj	fine	Underclass youth have those **[something]**, but not *those* ***opportunities***.

Table 2: Automatically generated coarse-grained and fine-grained positive interpretations

Step 2. Validating Potential Positive Interpretations

- *Given the negated statement, how much the statement [positive interpretation] below is true?*
- Answers: score from 0 to 5 (i.e., 0 - absolutely disagree, 5 - absolutely agree)

Example: You are not paying me for my overtime work.

Int. 1 [coarse]: **[some people]**'re paying me for my overtime work, but **not** ***you***.	0
Int. 2 [coarse]: You're paying **[somebody]** for my overtime work, but **not** ***me***.	1
Int. 3 [coarse]: You're paying me for **[something]**, but **not** ***for my overtime work***.	5
Int. 4 [fine]: You're paying me for **[somebody's]** overtime work, but not for **my** overtime work.	0
Int. 5 [fine]: You're paying me for my **[some adjective]** work, but **not for *my* overtime work.**	5
Int. 6 [fine]: You're paying me for my overtime **[something]**, but not for my **overtime work.**	0

Table 3: Positive interpretations and scores

Corpus Analysis

- 1,654 potential positive interpretations for verbal negation
- 781 from nominal negations
- 200 from adjectival negations
- Agreement: Inter-annotator Pearson correlation: 0.75
- At least one valid positive interpretation in 97% of negations on average

	#negs	%negs with≥1 valid PI	#PIs	Avg PIs per neg
Verbs	302	96.6	1654	5.50
Nouns	309	97.3	781	2.54
Adjs	75	98.0	200	2.67
Total	686	97.3	2635	3.57

Table 4: The total number of negations and the percentage with at least one valid positive interpretation, and the total number of positive interpretations generated and the average.

Learning to Score Potential Interpretations

- Standard supervised machine learning
  - Each potential positive interpretation along with their scores becomes an instance (1,700 instances)
  - 80 / 20 split (train / test)
    - ∗ all interpretations from a negation are either in the train or test split
- SVM for regression with RBF kernel
  - tuned using 10-fold cross validation, grid search

Results

Features	Gold	Predicted
neg_mark	-0.109	-0.077
basic	0.033	0.026
basic + path	0.474	0.482
basic + path + focus	0.530	0.560

Table 5: Results

Table 4 reports Pearson correlation for 4 different feature sets. Gold data set contains 379 test instances (20% of all annotations), however in the Predicted data some test instances are missing because the potential interpretations could not be generated.

Conclusions

- Humans intuitively understand negated statements in positive terms
- This paper presents a methodology to:
  - generate potential positive interpretations from verbal negation and
  - score them
- The procedure is grounded on syntactic dependencies

Forthcoming Research

We are increasing the number of negations and their probable positive interpretations, and apply a sequence to sequence deep learning models to generate them automatically.