Reference:

https://www.transacl.org/ojs/index.php/tacl/article/view/48

### Analysis and summary of

### “Good, Great, Excellent: Global Inference of Semantic Intensities”

### This paper attempts to create an intensity sale for adjectives using Mixed Integer Linear Programming (MILP). It tries to tackle the issue that standard information extraction methods do not provide enough context for real world decision making and the lack of evidence for many adjectival pairs.

### The authors’ main innovation is creating Weak-Strong (and Strong-Weak) ranking patterns, which imply ranking in pairs of words, and using it to bolster MILP. For example, *good* but not *great*, where but not is a Weak-Strong pattern. The authors generate pairwise scores by considering pairs of words in these patterns (both in the forward and backward direction and using both kinds of patterns). They then scan the corpus and collect frequencies of the patterns, as well as word frequencies. The generate a score that indicates the likelihood (or strength) of a word. A positive score for a pair x,y implies the x is a weaker term than y.

### Given a massive amount of pairwise scores, the authors then try to create a global information. MILP was chosen because it can be used to maximize the score under the constraints, where the constraints can be used to take advantage of pairwise scores. The implication being that how two adjectives relate to each other can imply something about the third. So scores aren’t treated as isolated instances. MILP weights pairs with higher scores more, indicating more confidence in their relative rankings.

**CORPUS:** Webscale, Google N-grams (1 to 5) with 1 trillion words and 95 billion sentences.

The testing was done on a set of annotated 569 word pairs and their results are compared to a Web baseline, a divide and conquer method (recursively splitting words until an order is formed), and a method devised by Sheinman and Tokunga using the synonymy information.