So-Called Non-Subsective Adjectives

Ellie Pavlick
University of Pennsylvania
epavlick@seas.upenn.edu

Chris Callison-Burch
University of Pennsylvania
ccb@cis.upenn.edu

Abstract

The interpretation of adjective-noun pairs plays a crucial role in tasks such as recognizing textual entailment. Formal semantics often places adjectives into a taxonomy which should dictate adjectives’ entailment behavior when placed in adjective-noun compounds. However, we show experimentally that the behavior of subsective adjectives (e.g. red) versus non-subsective adjectives (e.g. fake) is not as cut and dry as often assumed. For example, inferences are not always symmetric: while ID is generally considered to be mutually exclusive with fake ID, fake ID is considered to entail ID. We discuss the implications of these findings for automated natural language understanding.

1 Introduction

Most adjectives are subsective, meaning that an instance of an adjective-noun phrase is an instance of the noun: a red car is a car and a successful senator is a senator. In contrast, adjective-noun phrases involving non-subsective adjectives, such as imaginary and former (Table 1), denote a set that is disjoint from the denotation of the nouns they modify: an imaginary car is not a car and a former senator is not a senator. Understanding whether or not adjectives are subsective is critical in any task involving natural language inference. For example, consider the below sentence pair from the Recognizing Textual Entailment (RTE) task (Giampiccolo et al., 2007):

1. (a) U.S. District Judge Leonie Brinkema accepted would-be hijacker Zacarias Moussaoui’s guilty pleas . . .
   (b) Moussaoui participated in the Sept. 11 attacks.

In this example, recognizing that 1(a) does not entail 1(b) hinges on understanding that a would-be hijacker is not a hijacker.

The observation that adjective-nouns (ANs) involving non-subsective adjectives do not entail the underlying nouns (Ns) has led to the generalization that the deletion of non-subsective adjectives tends to result in contradictory utterances: Moussaoui is a would-be hijacker entails that it is not the case that Moussaoui is a hijacker. This generalization has prompted normative rules for the treatment of such adjectives in various NLP tasks. In information extraction, it is assumed that systems cannot extract useful rules from sentences containing non-subsective modifiers (Angeli et al., 2015), and in RTE, it is assumed that systems should uniformly penalize insertions and deletions of non-subsective adjectives (Amoia and Gardent, 2006).

While these generalizations are intuitive, there is little experimental evidence to support them. In this paper, we collect human judgements of the validity of inferences following from the insertion and deletion of various classes of adjectives and analyze the results. Our findings suggest that, in practice, most sentences involving non-subsective ANs can be safely generalized to statements about the N. That is, non-subsective adjectives often behave like normal, subsective adjectives. On further analysis, we reveal that, when adjectives do behave non-subsectively, they often exhibit asymmetric entailment behavior in which inser-

<table>
<thead>
<tr>
<th>Privative Non-Subsective (AN ∩ N = ∅)</th>
<th>Plain Non-Subsective (AN ⊄ N and AN ∩ N ≠ ∅)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti- artificial counterfeit deputy</td>
<td>alleged apparent arguable assumed doubtful</td>
</tr>
<tr>
<td>erstwhile ex- fabricated fake</td>
<td>believed debatable disputed doubtful</td>
</tr>
<tr>
<td>false fictional fictitious former</td>
<td>dubious erroneous expected faulty</td>
</tr>
<tr>
<td>hypothetical imaginary mock mythical</td>
<td>future historic impossible improbable</td>
</tr>
<tr>
<td>onetime past phony pseudo-</td>
<td>likely mistaken ostensible plausible</td>
</tr>
<tr>
<td>simulated spurious virtual would-be</td>
<td>possible potential predicted presumed</td>
</tr>
<tr>
<td></td>
<td>probable proposed putative questionable</td>
</tr>
<tr>
<td></td>
<td>seeming so-called supposed suspicious</td>
</tr>
<tr>
<td></td>
<td>theoretical uncertain unlikely unsuccessful</td>
</tr>
</tbody>
</table>

Table 1: 60 non-subsective adjectives from Nayak et al. (2014). Noun phrases involving non-subsective adjectives are assumed not to entail the head noun. E.g. would-be hijacker ≠ hijacker. (See Section 2 for definition of privative vs. plain).
tion leads to contradictions \((\text{ID} \Rightarrow \neg \text{fake ID})\) but deletion leads to entailments \((\text{fake ID} \Rightarrow \text{ID})\). We present anecdotal evidence for how the entailment associated with inserting/deleting a non-subsective adjective depends on the salient properties of the noun phrase under discussion, rather than on the adjective itself.

2 Background and Related Work

Classes of Adjectives. Adjectives are commonly classified taxonomically as either subsective or non-subsective (Kamp and Partee, 1995). Subsective adjectives are adjectives which pick out a subset of the set denoted by the unmodified noun; that is, \(\text{AN} \subseteq \text{N}\). For non-subsective adjectives, in contrast, the AN cannot be guaranteed to be a subset of N. For example, \text{clever} is subsective, and so a \text{clever thief} is always a \text{thief}. However, \text{alleged} is non-subsective, so there are many possible worlds in which an \text{alleged thief} is not in fact a \text{thief}. Of course, there may also be many possible worlds in which the \text{alleged thief} is a \text{thief}, but the word \text{alleged}, being non-subsective, does not guarantee this to hold.

Non-subsective adjectives can be further divided into two classes: privative and plain. Sets denoted by privative ANs are completely disjoint from the set denoted by the head N (\(\text{AN} \cap \text{N} = \emptyset\)), and this mutual exclusivity is encoded in the meaning of the A itself. For example, \text{fake} is considered to be a quintessential privative adjective since, given the usual definition of \text{fake}, a \text{fake ID} cannot actually be an \text{ID}. For plain non-subsective adjectives, there may be worlds in which the AN is and N, and worlds in which the AN is not an N: neither inference is guaranteed by the meaning of the A. As mentioned above, \text{alleged} is quintessentially plain non-subsective since, for example, an \text{alleged thief} may or may not be an actual \text{thief}. In short, we can summarize the classes of adjectives in the following way: subsective adjectives entail the nouns they modify, privative adjectives contradict the nouns they modify, and plain non-subsective adjectives are compatible with (but do not entail) the nouns they modify. Figure 1 depicts these distinctions.

While the hierarchical classification of adjectives described above is widely accepted and often applied in NLP tasks (Amoia and Gardent, 2006; Amoia and Gardent, 2007; Boleda et al., 2012; McCrae et al., 2014), it is not undisputed. Some linguists take the position that in fact privative adjectives are simply another type of subsective adjective (Partee, 2003; McNally and Boleda, 2004; Abdullah and Frost, 2005; Partee, 2007). Advocates of this theory argue that the denotation of the noun should be expanded to include both the properties captured by the privative adjectives as well as those captured by the subsective adjectives. This expanded denotation can explain the acceptability of the sentence \text{Is that gun real or fake?}, which is difficult to analyze if \text{gun} entails \neg \text{fake gun}. More recent theoretical work argues that common nouns have a “dual semantic structure” and that non-subsective adjectives modify part of this meaning (e.g. the functional features of the noun) without modifying the extension of the noun (Del Pinal, 2015). Such an analysis can explain how we can understand a \text{fake gun} as having many, but not all, of the properties of a \text{gun}.

Several other studies abandon the attempt to organize adjectives taxonomically, and instead focus on the properties of the modified noun. Nayak et al. (2014) categorize non-subsective adjectives in terms of the proportion of properties that are shared between the N and the AN and Pustejovsky (2013) focus on syntactic cues about exactly which properties are shared. Bakhshandh and Allen (2015) analyze adjectives by observing that, e.g., \text{red} modifies \text{color} while \text{tall} modifies \text{size}. In Section 5, we discuss the potential benefits of pursuing these property-based analyses in relation to our experimental findings.

---

1We use the notation N and AN to refer both the the natural language expression itself (e.g. \text{red car}) as well as its denotation, c.g. \(\{x|\text{x is a red car}\}\).
Recognizing Textual Entailment. We analyze adjectives within the context of the task of Recognizing Textual Entailment (RTE) (Dagan et al., 2006). The RTE task is defined as: given two natural language utterances, a premise $p$ and a hypothesis $h$, would a typical human reading $p$ likely conclude that $h$ is true? We consider the RTE task as a three-way classification: ENTAILMENT, CONTRADICTION, or UNKNOWN (meaning $p$ neither entails nor contradicts $h$).

3 Experimental Design

Our goal is to analyze how non-subjective adjectives effect the inferences that can be made about natural language. We begin with the set of 60 non-subjective adjectives identified by Nayak et al. (2014), which we split into plain non-subjective and privative adjectives (Table 1). We search through the Annotated Gigaword corpus (Napoles et al., 2012) for occurrences of each adjective in the list, restricting to cases in which the adjective appears as an adjective modifier of (is in an amod dependency relation with) a common noun (NN). For each adjective, we choose 10 sentences such that the adjective modifies a different noun in each. As a control, we take a small sample 100 ANs chosen randomly from our corpus. We expect these to contain almost entirely subjective adjectives.

For each selected sentence $s$, we generate $s'$ by deleting the non-subjective adjective from $s$. We then construct two RTE problems, one in which $p = s$ and $h = s'$ (the deletion direction), and one in which $p = s'$ and $h = s$ (the insertion direction). For each RTE problem, we ask annotators to indicate on a 5-point scale how likely it is that $p$ entails $h$, where a score of -2 indicates definite contradiction and a score of 2 indicates definite entailment. We use Amazon Mechanical Turk, requiring annotators to pass a qualification test of simulating two RTE problems, one in which a premise $p$ and a hypothesis $h$, would a typical human reading $p$ likely conclude that $h$ is true? We consider the RTE task as a three-way classification: ENTAILMENT, CONTRADICTION, or UNKNOWN (meaning $p$ neither entails nor contradicts $h$).

4 Results

Expectations. Based on the theoretical adjective classes described in Section 2, we expect that both the insertion and the deletion of privative adjectives from a sentence should result in judgments of CONTRADICTION: i.e. it should be the case that $\text{fake ID} \Rightarrow \sim \text{ID}$ and $\text{ID} \Rightarrow \sim \text{fake ID}$. Similarly, we expect plain non-subjective adjectives to receive labels of UNKNOWN in both directions. We expect the subjective adjectives to receive labels of ENTAILMENT in the deletion direction (red car ⇒ car) and labels of UNKNOWN in the insertion direction (car ≠ red car). Figure 1 depicts these expected distributions.

Observations. The observed entailment patterns for insertion and deletion of non-subjective adjectives are shown in Figure 2. Our control sample of subjective adjectives (Figure 2c) largely produced the expected results, with 96% of deletions producing ENTAILMENTS and 73% of insertions producing UNKNOWNS. The entailment patterns produced by the non-subjective adjectives, however, did not match our predictions. The plain non-subjective adjectives (e.g. alleged) behave nearly identically to how we expect regular, subjective adjectives to behave (Figure 2b). That is, in 80% of cases, deleting the plain non-subjective adjective was judged to produce ENTAILMENT, rather than the expected UNKNOWN. The examples in Table 2 shed some light onto why this is the case. Often, the differences between N and AN are not relevant to the main point of the utterance. For example, while an expected surge in unemployment is not a surge in unemployment, a policy that deals with an expected surge deals with a surge.

| (1) | Swiss officials on Friday said they’ve launched an investigation into Urs Tinner’s alleged role. |
| (2) | To deal with an expected surge in unemployment, the plan includes a huge temporary jobs program. |
| (3) | They kept it close for a half and had a theoretical chance come the third quarter. |

Table 2: Contrary to expectations, the deletion of plain non-subjective adjectives often preserves the (plausible) truth in a model. E.g. alleged role ≁ role, but investigation into alleged role ⇒ investigation into role.

The privative adjectives (e.g. fake) also fail to match the predicted distribution. While insertions often produce the expected CONTRADICTIONS, deletions produce a surprising number of ENTAILMENTS (Figure 2a). Such a pattern does not fit into any of the adjective classes from Figure 1. While some ANs (e.g. counterfeit money) behave in the prototypically privative way, others (e.g. mythical beast) have the property in which N ⇒ ¬AN, but AN ⇒ N (Figure 3). Table 3 provides some telling examples of how this AN⇒N inference,
Figure 2: Observed entailment judgements for insertion (blue) and deletion (red) of adjectives. Compare to expected distributions in Figure 1.

(1a) ENTAIL. Flawed counterfeit software can corrupt the information entrusted to it.
(1b) CONTRA. Pharmacists in Algodones denied selling counterfeit medicine in their stores.
(2a) ENTAIL. He also took part in a mock debate Sunday.
(2b) CONTRA. Investigation leader said the prisoner had been subjected to a mock execution.
(3a) ENTAIL. The plants were grown under artificial light and the whole operation was computerised.
(3b) UNKNOWN Thrun predicted that leaps in artificial intelligence would lead to driverless cars on the roads by 2030.

Table 3: Entailment judgements for the deletion of various privative adjectives from a sentence. Whether or not deletion results in CONTRADICTION depends on which properties of the noun are most relevant.

in the case of privative adjectives, often depends less on the adjective itself, and more on properties of the modified noun that are at issue in the given context. For example, in Table 3 Example 2(a), a mock debate probably contains enough of the relevant properties (namely, arguments) that it can entail debate, while in Example 2(b), a mock execution lacks the single most important property (the death of the executee) and so cannot entail execution. (Note that, from Example 3(b), it appears the jury is still out on whether leaps in artificial intelligence entail leaps in intelligence...)

5 Discussion

The results presented suggest a few important patterns for NLP systems. First, that while a non-subsective AN might not be an instance of the N (taxonomically speaking), statements that are true of an AN are often true of the N as well. This is relevant for IE and QA systems, and is likely to become more important as NLP systems focus more on “micro reading” tasks (Nakashole and Mitchell, 2014), where facts must be inferred from single documents or sentences, rather than by exploiting the massive redundancy of the web. Second, the asymmetric entailments associated with privative adjectives suggests that the contradictions generated by privative adjectives may not be due to a strict denotational contradiction, but rather based on implicature: i.e. if an ID is in fact fake, the speaker is obligated to say so, and thus, when ID appears unmodified, it is fair to assume it is not a fake ID. Testing this hypothesis is left for future research. Finally, the examples in Tables 2 and 3 seem to favor a properties-oriented analysis of adjective semantics, rather than the taxonomic analysis often used. Nayak et al. (2014)’s attempt to characterize adjectives in terms of the number of properties the AN shares with N is a step in the right direction, but it seems that what is relevant is not how many properties are shared, but rather which properties are shared, and which properties are at issue in the given context.

6 Conclusion

We present experimental results on textual inferences involving non-subsective adjectives. We show that, contrary to expectations, the deletion of non-subsective adjectives from a sentence does not necessarily result in
non-entailment. Thus, in applications such as information extraction, it is often possible to extract true facts about the N from sentences involving a non-subjective AN. Our data suggests that inferences involving non-subjective adjectives require more than strict reasoning about denotations, and that a treatment of non-subjective adjectives based on the properties of the AN, rather than its taxonomic relation to the N, is likely to yield useful insights.

Acknowledgments

This research was supported by a Facebook Fellowship, and by gifts from the Alfred P. Sloan Foundation, Google, and Facebook. This material is based in part on research sponsored by the NSF grant under IIS-1249516 and DARPA under number FA8750-13-2-0017 (the DEFT program). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes. The views and conclusions contained in this publication are those of the authors and should not be interpreted as representing official policies or endorsements of DARPA and the U.S. Government.

We would like to thank the anonymous reviewers for their very thoughtful comments. We would also like to thank the Mechanical Turk annotators for their contributions.

References


