Hypernyms as Answer Types

Gregory Marton GREMIO@CSAIL.MIT.EDU
Stefanie Tellex STEFIE10@ALUM.MIT.EDU
Aaron Fernandes ADFERNAN@ALUM.MIT.EDU
Boris Katz BORIS@CSAIL.MIT.EDU
MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge MA, 02139 USA

1. Introduction

A system that answers questions automatically with just the right information will return answers of the correct type. The question itself often specifies the answer type it expects. When answering a question like "*What flowers did Van Gogh paint?*", we prefer answers that are members of the class of flowers, e.g. tulips, begonias, or sunflowers, over other things Van Gogh might paint, like "watercolors".

The answer type is also a poor search term: for the question "What countries produce coffee?", answers may be of the form "Brazil produces coffee", far from any mention of the term "country".

2. Related Work

The first Answer Type Identification systems used named entity categories to require, for example, `person` or `organization` types for "Who" questions, and `date` or `time` types for "When" questions (Hirschman & Gaizauskas, 2001).

Today's top systems use answer type taxonomies with hundreds to thousands of entries (Voorhees, 2004; Hovy et al., 2001; Hovy et al., 2002; Katz et al., 2003; Katz et al., 2004), allowing systems to answer questions that, for example, start with "Which countries" or "What president".

Separately, others have explored automatic identification of hypernym–hyponym (henceforth "class"–"member") relations in large bodies of text (Hearst, 1992; Carballo, 1999; Gruenstein, 2001; Snow et al., ; Pantel, 2005). Phrases in English like "corn, wheat, and other staple crops" identify corn and wheat as members of the class of "staple crops". For the most part, these techniques have been used to augment WordNet, a machine–readable lexicon of English that encodes, among other things, these class–member relations.

Harabagiu *et al.* (Pasca & Harabagiu, 2001) observed that answer types can be identified with WordNet classes, and their members can be used as a set of possible answers. Fleischman (Fleischman et al., 2003) used automatic class identification to answer "Who is *P*?" with the automatically extracted classes of which *P* is a member. We are the first to put together these two ideas: to automatically identify classes in the same text that we will answer questions on, and to use those classes as answer types. The class information we extract also turns out to be instrumental in a number of related question–answering tasks.

3. Approach

We build on the observation that classes are answer types and their members are good candidate answers, and we use automatic class identification techniques to identify over a million classes and their members. These enable us to accurately answer questions with much more specific answer types than those previously available. Examples include: "Which non–OPEC countries ..." and "Which former Yugoslav President ... ".

3.1 Candidate Classes

Aaron Fernandes extended (Fernandes, 2004) Fleischman *et al.*'s work (Fleischman et al., 2003) on finding definitions from applying only to person names, to applying to most noun phrases. He generated candidate class–member pairs like those in Figure 1.

3.2 Aggregating Classes

Marton and Tellex then aggregated these class–member candidate pairs π, using a probability of correctness $\mathrm{Freq}(\pi)$ based on the number of times a pair occurred in the corpus, and on the precision $p(z)$ of each pattern z: [1]

$$\mathrm{Freq}(\pi) = \frac{\sum_z p(z) * \mathrm{count}(\pi)}{\sum_z (p(z) * \sum_i \mathrm{count}(i))}$$

[1] The numerator is an expectation of times π was correctly seen; i in the denominator iterates over all pairs observed with the pattern z, making the denominator a normalizing constant. $p(z)$ was estimated for each z from at least 100 examples.

Pattern cue	$p(z)$	Frequency	Example
common noun then name	0.75	2,125,812	*President* `Clinton`
apposition marked by commas	0.89	625,962	`Noemi Sanin`, *a former foreign minister*,
plural then like	0.42	158,167	*immunisable diseases* like `polio`.
such as or such *X* as	0.47	118,684	*stinging insects*, such as `bees`, `wasps`, `hornets` and `red ants`,
called or also called	0.70	14,066	*a game* called `Tightrope walker`
named then proper name	0.64	8,992	*an Armenian* named `Wilhelm Vigen`
known as, also known as	0.68	8,199	*low-tariff trade rights* known as `most-favored-nation status`

Figure 1. Examples of the seven patterns expressing class–member relations. Precision (*p(z)*) and Frequency of each pattern are shown for a million–article body of newspaper text, along with an example *class* and `member` in context.

4. Progress

We have collected class–member pairs from the AQUAINT corpus[2] of English newspaper text using the patterns described above, yielding 2.3 million candidate categories. Precision, measured using human correctness judgements, is around 50%. We measured recall using the NIST TREC 2003 and 2004 Question Answering Track[3] data. The first form of recall that we measured tested the coverage of categories: of the focus–phrases of questions asked ("Which *flowers* did Van Gogh paint"), around 91% were candidate categories in our list (95% for 2005). The second form of recall tested the coverage of members: for each question, about 30% of the known answers were members of a class associated with the question's focus phrase.[4]

5. Future Work

We will integrate this work with our question–answering system in a number of ways:

- If the answer type in a question (between the *Wh*–word and the verb) matches a class, then we will use members of that class as search terms, and we will prefer those members as answers to the question.

- Categories may appear elsewhere in the question, as in the question: "Which *oil companies* drill in *non–OPEC countries*?". When we identify such categories, we will again use members as search terms.

- In a conversation, a question might refer to a contextual topic by one of its classes. For example, if "Conde Nast" is under discussion, one must know that it is a publishing company in order to understand the question: "Who is *the publishing company*'s CEO?"

[2]http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002T31

[3]http://trec.nist.gov/

[4]Answers for 2005 will not become available until November.

Automatic Hypernym Extraction and Automatic Question Answering have been areas of intense study since the 1990s. Question answering is most often applied to a particular body of text. This work lets an automatic system read that text to learn the class information it needs to answer questions about the knowledge within.

References

Carballo, S. (1999). Automatic construction of a hypernym-labeled noun hierarchy from text. *ACL1999*.

Fernandes, A. (2004). Answering definitional questions before they are asked. Master's thesis, MIT.

Fleischman, M., Hovy, E., & Echihabi, A. (2003). Offline strategies for online question answering: Answering questions before they are asked. *ACL2003*.

Gruenstein, A. (2001). Learning hypernyms from corpora.

Hearst, M. (1992). Automatic acquisition of hyponyms from large text corpora. *14th COLING conference*.

Hirschman, L., & Gaizauskas, R. (2001). Natural language question answering: The view from here. *Natural Language Engineering*.

Hovy, E., Hermjakob, U., & Lin, C.-Y. (2001). The use of external knowledge in factoid QA. *TREC-10*.

Hovy, E., Hermjakob, U., & Ravichandran, D. (2002). A question/answer typology with surface text patterns. *HLT2002*.

Katz, B., et al. (2003). Integrating web-based and corpus-based techniques for question answering. *TREC-12*.

Katz, B., et al. (2004). Answering multiple questions on a topic from heterogeneous resources. *TREC-13*.

Pantel, P. (2005). Inducing ontological co-occurrence vectors. *ACL2003*.

Pasca, M., & Harabagiu, S. (2001). The informative role of WordNet in open-domain question answering. *NAACL 2001 Workshop on WordNet and Other Lexical Resources* (pp. 138–143).

Snow, R., Jurafsky, D., & Ng, A. Y. Learning syntactic patterns for automatic hypernym discovery. *NIPS2004*.

Voorhees, E. (2004). Overview of the TREC 2004 question answering track.