Table 2: Example non-coreferent paths: *Italicized* entities do *not* generally corefer

	Pattern	Example
1.	*Noun* thanked ... for *pronoun*'s assistance	*John* thanked him for *his* assistance.
2.	*Noun* wanted *pronoun* to lie.	*The president* wanted *her* to lie.
3.	... *Noun* into *pronoun*'s pool	Max put the *floaties* into *their* pool.
4.	... use *Noun* to *pronoun*'s advantage	The company used *the delay* to *its* advantage.
5.	*Noun* suspended *pronoun*	*Mary* suspended *her*.
6.	*Noun* was *pronoun*'s relative.	*The Smiths* were *their* relatives.
7.	*Noun* met *pronoun*'s demands	*The players' association* met *its* demands.
8.	... put *Noun* at the top of *pronoun*'s list.	The government put *safety* at the top of *its* list.

tem. We also tried expanding our coverage by using paths *similar* to paths with known path coreference (based on distributionally similar words), but this did not generally increase performance.

4 Bootstrapping in Pronoun Resolution

Our determination of path coreference can be considered a bootstrapping procedure. Furthermore, the coreferent paths themselves can serve as the seed for bootstrapping additional coreference information. In this section, we sketch previous approaches to bootstrapping in coreference resolution and explain our new ideas.

Coreference bootstrapping works by assuming resolutions in unlabelled text, acquiring information from the putative resolutions, and then making inferences from the aggregate statistical data. For example, we assumed two pronouns from the same pronoun group were coreferent, and deduced path coreference from the accumulated counts.

The potential of the bootstrapping approach can best be appreciated by imagining millions of documents with coreference annotations. With such a set, we could extract fine-grained features, perhaps tied to individual words or paths. For example, we could estimate the likelihood each noun belongs to a particular gender/number class by the proportion of times this noun was labelled as the antecedent for a pronoun of this particular gender/number.

Since no such corpus exists, researchers have used coarser features learned from smaller sets through supervised learning (Soon et al., 2001; Ng and Cardie, 2002), manually-defined coreference patterns to mine specific kinds of data (Bean and Riloff, 2004; Bergsma, 2005), or accepted the noise inherent in unsupervised schemes (Ge et al., 1998; Cherry and Bergsma, 2005).

We address the drawbacks of these approaches by using coreferent paths as the assumed resolutions in the bootstrapping. Because we can vary the threshold for defining a coreferent path, we can trade-off coverage for precision. We now outline two potential uses of bootstrapping with coreferent paths: learning gender/number information (Section 4.1) and augmenting a semantic compatibility model (Section 4.2). We bootstrap this data on our automatically-parsed news corpus. The corpus comprises 85 GB of news articles taken from the world wide web over a 1-year period.

Table 3: Gender classification performance (%)

Classifier	F-Score
Bergsma (2005) Corpus-based	85.4
Bergsma (2005) Web-based	90.4
Bergsma (2005) Combined	92.2
Duplicated Corpus-based	88.0
Coreferent Path-based	90.3

4.1 Probabilistic Gender/Number

Bergsma (2005) learns noun gender (and number) from two principal sources: 1) mining it from manually-defined lexico-syntactic patterns in parsed corpora, and 2) acquiring it on the fly by counting the number of pages returned for various gender-indicating patterns by the Google search engine. The web-based approach outperformed the corpus-based approach, while a system that combined the two sets of information resulted in the highest performance (Table 3). The combined gender-classifying system is a machine-learned classifier with 20 features.

The time delay of using an Internet search engine within a large-scale anaphora resolution effort is currently impractical. Thus we attempted