A Rule-Based System for Unrestricted Bridging Resolution: Recognizing Bridging Anaphora and Finding Links to Antecedents

Yufang Hou[1], Katja Markert[2], Michael Strube[1]
[1] Heidelberg Institute for Theoretical Studies gGmbH, Heidelberg, Germany
(yufang.hou|michael.strube)@h-its.org
[2]School of Computing, University of Leeds, UK
scskm@leeds.ac.uk

Abstract

Bridging resolution plays an important role in establishing (local) entity coherence. This paper proposes a rule-based approach for the challenging task of unrestricted bridging resolution, where bridging anaphors are not limited to definite NPs and semantic relations between anaphors and their antecedents are not restricted to meronymic relations. The system consists of eight rules which target different relations based on linguistic insights. Our rule-based system significantly outperforms a reimplementation of a previous rule-based system (Vieira and Poesio, 2000). Furthermore, it performs better than a learning-based approach which has access to the same knowledge resources as the rule-based system. Additionally, incorporating the rules and more features into the learning-based system yields a minor improvement over the rule-based system.

1   Introduction

Bridging resolution recovers the various non-identity relations between anaphora and antecedents. It plays an important role in establishing entity coherence in a text. In Example 1, the links between the bridging anaphors (**The five astronauts** and **touchdown**) and the antecedent (*The space shuttle Atlantis*) establish (local) entity coherence.[1]

(1)   *The space shuttle Atlantis* landed at a desert air strip at Edwards Air Force Base, Calif., ending a five-day mission that dispatched the Jupiter-bound Galileo space probe. **The**

five astronauts returned to Earth about three hours early because high winds had been predicted at the landing site. Fog shrouded the base before **touchdown**.

Bridging or *associative anaphora* has been widely discussed in the linguistic literature (Clark, 1975; Prince, 1981; Gundel et al., 1993; Löbner, 1998). Poesio and Vieira (1998) and Bunescu (2003) include cases where antecedent and anaphor are coreferent but do not share the same head noun (*different-head coreference*). We follow our previous work (Hou et al., 2013b) and restrict *bridging* to non-coreferential cases. We also exclude *comparative anaphora* (Modjeska et al., 2003).

Bridging resolution includes two subtasks: (1) recognizing bridging anaphors and (2) finding the correct antecedent among candidates. In recent empirical work, these two subtasks have been tackled separately: (Markert et al., 2012; Cahill and Riester, 2012; Rahman and Ng, 2012; Hou et al., 2013a) handle bridging recognition as part of information status (IS) classification, while (Poesio et al., 1997; Poesio et al., 2004; Markert et al., 2003; Lassalle and Denis, 2011; Hou et al., 2013b) concentrate on antecedent selection only, assuming that bridging recognition has already been performed. One exception is Vieira and Poesio (2000). They propose a rule-based system for processing definite NPs. However, they include *different-head coreference* into *bridging*. They report results for the whole anaphora resolution but do not report results for bridging resolution only. Another exception is Rösiger and Teufel (2014). They apply a coreference resolution system with several additional semantic features to find bridging links in scientific text where bridging anaphors are limited to definite NPs. They report preliminary results using the CoNLL scorer. However, we think the coreference resolution system and the evaluation metric for coreference resolution are

---

[1]Examples are from OntoNotes (Weischedel et al., 2011). Bridging anaphora are typed in boldface; antecedents in italics.

not suitable for bridging resolution since bridging is not a set problem.

Another vein of research for bridging resolution focuses on formal semantics. Asher and Lascarides (1998) and Cimiano (2006) model bridging by integrating discourse structure and semantics from a formal semantics viewpoint. However, the implementation of such a theoretical framework is beyond the current capabilities of NLP since it depends heavily on commonsense entailment.

In this paper, we propose a rule-based system for unrestricted bridging resolution. The system consists of eight rules which we carefully design based on linguistic intuitions, i.e., how the nature of bridging is reflected by various lexical, syntactic and semantic features. We evaluate our rule-based system on a corpus where bridging is reliably annotated. We find that our rule-based system significantly outperforms a reimplementation of a previous rule-based system (Vieira and Poesio, 2000). We further notice that our rule-based system performs better than a learning-based approach which has access to the same knowledge resources as the rule-based system. Surprisingly, incorporating the rules and more features into the learning-based approach only yields a minor improvement over the rule-based system. We observe that diverse bridging relations and relatively small-scale data for each type of relations make generalization difficult for the learning-based approach. This work is – to the best of our knowledge – the first system recognizing bridging anaphora and finding links to antecedents for unrestricted phenomenon where bridging anaphors are not limited to definite NPs and semantic relations between anaphors and their antecedents are not restricted to meronymic relations.

2 Data

All the data used throughout the paper come from the ISNotes corpus[2] released by Hou et al. (2013b). This corpus contains around 11,000 NPs annotated for information status including 663 bridging NPs and their antecedents in 50 texts taken from the WSJ portion of the OntoNotes corpus (Weischedel et al., 2011). ISNotes is reliably annotated for bridging: for bridging anaphor recognition, κ is over 60 for all three possible an-

notator pairings (κ is over 70 for two expert annotators); for selecting bridging antecedents, agreement is around 80% for all annotator pairings.

It is notable that bridging anaphors in ISNotes are not limited to definite NPs as in previous work (Poesio et al., 1997; Poesio et al., 2004; Lassalle and Denis, 2011). Table 1 shows the bridging

Bridging Anaphors	663
Non-determiner	44.9%
Definite	38.5%
Indefinite	15.4%
Other-determiner	1.2%

Table 1: Bridging anaphora distribution w.r.t. determiners in ISNotes.

anaphora distribution with regard to determiners in ISNotes: only around 38% of bridging anaphors are definite NPs (NPs modified by *the*); 15.4% of bridging anaphors are modified by determiners such as *a*, *an* or *one* which normally indicate indefinite NPs. Most bridging anaphors (43%) are not modified by any determiners, such as **touchdown** in Example 1. A small fraction of bridging anaphors (1.2%) are modified by other determiners, such as demonstratives.

The semantic relations between anaphor and antecedent in the corpus are extremely diverse: only 14% of anaphors have a part-of/attribute-of relation with the antecedent (see Example 2) and only 7% of anaphors stand in a set relationship to the antecedent (see Example 3). 79% of anaphors have "other" relations with their antecedents (without further distinction), including encyclopedic relations such as *The space shuttle Atlantis*-**The five astronauts** (see Example 1) as well as context-specific relations such as *The space shuttle Atlantis*-**touchdown** (Example 1).

(2) At age eight, Josephine Baker was sent by her mother to *a white women's house* to do chores in exchange for meals and a place to sleep – a place in **the basement** with coal.

(3) This creates *several problems*. **One** is that there are not enough police to satisfy small businesses.

In ISNotes, bridging anaphora with distant antecedents are common when the antecedent is the global focus of a document. 29% of the anaphors in the corpus have antecedents that are three or more sentences away.

Bridging resolution is an extremely challenging task in ISNotes. In contrast with surface clues for coreference resolution, there are no clear surface clues for bridging resolution. In Example 4, the bridging anaphor **low-interest disaster loans** associates to the antecedent *the Carolinas and Caribbean*, whereas in Example 5 the NP loans is a generic use. In Example 6, the bridging anaphor **The opening show** associates to the antecedent *Mancuso FBI*, whereas the NP the show is coreferent with its antecedent *Mancuso FBI*.

(4) The $2.85 billion measure comes on top of $1.1 billion appropriated after Hugo stuck *the Carolinas and Caribbean* last month, and these totals don't reflect the additional benefit of **low-interest disaster loans**.

(5) Many states already have Enterprise Zones and legislation that combines tax incentives, loans, and grants to encourage investment in depressed areas.

(6) Over the first few weeks, *Mancuso FBI* has sprung straight from the headlines. **The opening show** featured a secretary of defense designate accused of womanizing (a la John Tower).
. . .
Most of all though, the show is redeemed by the character of Mancuso.

Our previous work on bridging resolution on this corpus only focuses on its subtasks. In Hou et al. (2013a) we model bridging anaphora recognition as a subtask of learning fine-grained information status. We report an F-measure of 0.42 for bridging anaphora recognition. In Hou et al. (2013b) we propose a joint inference framework for antecedent selection by exploring Markov logic networks. We report an accuracy of 0.41 for antecedent selection given gold bridging anaphora. In this paper, we aim to solve these two substasks together, i.e., recognizing bridging anaphora and finding links to antecedents.

3 Method

In this section, we describe our rule-based system for unrestricted bridging resolution. We choose ten documents randomly from the corpus as the development set. Then we carefully design rules for finding "bridging links" among all NPs in a

document based on the generalizations of bridging in the linguistic literature as well as our inspections of bridging annotations in the development set. The system consists of two components: bridging link prediction and post processing.

3.1 Bridging Link Prediction

The bridging link prediction component consists of eight rules. Löbner (1985; 1998) interprets bridging anaphora as a particular kind of functional concept, which in a given situation assign a necessarily unique correlate to a (implicit) possessor argument. He distinguishes between relational nouns (e.g. parts terms, kinship terms, role terms) and sortal nouns and points out that relational nouns are more frequently used as bridging anaphora than sortal nouns. Rule1 to Rule4 in our system aim to resolve such relational nouns. We design Rule5 and Rule6 to capture set bridging. Finally, Rule7 and Rule8 are motivated by previous work on implicit semantic role labeling (Laparra and Rigau, 2013) which focuses on few predicates.

For all NPs in a document, each rule r is applied separately to predict a set of potential bridging links. Every rule has its own constraints on bridging anaphora and antecedents respectively. Bridging anaphors are diverse with regard to syntactic form and function: they can be modified by definite or indefinite determiners (Table 1), furthermore they can take the subject (e.g. Example 3 and Example 6) or other positions (e.g. Example 2 and Example 4) in sentences. The only frequent syntactic property shared is that bridging anaphors most often have a simple internal structure concerning modification. Therefore we first create an initial list of potential bridging anaphora A which excludes NPs which have a complex syntactic structure. An NP is added to A if it does not contain any other NPs and do not have modifications strongly indicating comparative NPs (such as *other symptoms*)[3]. Since head match is a strong indicator of coreference anaphora for definite NPs (Vieira and Poesio, 2000; Soon et al., 2001), we further exclude definite NPs from A if they have the same head as a previous NP. Then a set of potential bridging anaphors A_r is chosen from A based on r's constraints on bridging anaphora. Finally, for each potential bridging anaphor $ana \in$

---

[3]A small list of 10 markers such as *such, another . . .* and the presence of adjectives or adverbs in the comparative form are used to predict comparative NPs.

A_r, a single best antecedent *ante* from a list of candidate NPs (C_{ana}) is chosen if the rule's constraint on antecedents is applied successfully.

Every rule has its own scope to form the antecedent candidate set C_{ana}. Instead of using a static sentence window to construct the list of antecedent candidates like most previous work for resolving bridging anaphora (Poesio et al., 1997; Markert et al., 2003; Poesio et al., 2004; Lassalle and Denis, 2011), we use the development set to estimate the proper scope for each rule. The scope is influenced by the following factors: (1) the nature of the target bridging link (e.g., set bridging is a local coherence phenomenon where the antecedent often occurs in the same or up to two sentences prior to the anaphor); and (2) the strength of the rule's constraint to select the correct antecedent (e.g., in Rule8, the ability to select the correct antecedent decreases with increasing the scope to contain more antecedent candidates). In the following, we describe the motivation for each rule and their constraints in detail.

Rule1: building part NPs. To capture typical part-of bridging (see Example 2), we extract a list of 45 nouns which specify building parts (e.g. *room* or *roof*) from the General Inquirer lexicon (Stone et al., 1966). A common noun phrase from A is added to A_{r1} if: (1) its head appears in the building part list; and (2) it does not contain any nominal pre-modifications. Then for each potential bridging anaphor $ana \in A_{r1}$, the NP with the strongest semantic connectivity to the potential anaphor ana among all NPs preceding ana from the same sentence as well as from the previous two sentences is predicted to be the antecedent.

The semantic connectivity of an NP to a potential anaphor is measured via the hit counts of the preposition pattern query (*anaphor preposition NP*) in big corpora[4]. An initial effort to extract partOf relations using WordNet yields low recall on the development set. Therefore we use semantic connectivity expressed by prepositional patterns (e.g. *the basement of the house*) to capture underlying semantic relations. Such syntactic patterns are also explored in Poesio et al. (2004) to resolve meronymy bridging.

---

[4]We use Gigaword (Parker et al., 2011) with automatic POS tag and NP chunk information.

Rule2: relative person NPs. This rule is used to capture the bridging relation between a relative (e.g. *The husband*) and its antecedent (e.g. *She*). A list of 110 such relative nouns is extracted from WordNet. However, some relative nouns are frequently used generically instead of being bridging, such as *children*. To exclude such cases, we compute the argument taking ratio α for an NP using NomBank (Meyers et al., 2004). For each NP, α is calculated via its head frequency in the NomBank annotation divided by the head's total frequency in the WSJ corpus in which the NomBank annotation is conducted. The value of α reflects how likely an NP is to take arguments. For instance, the value of α is 0.90 for *husband* but 0.31 for *children*. To predict bridging anaphora more accurately, a conservative constraint is used. An NP from A is added to A_{r2} if: (1) its head appears in the relative person list; (2) its argument taking ratio α is bigger than 0.5; and (3) it does not contain any nominal or adjective pre-modifications. Then for each potential bridging anaphor $ana \in A_{r2}$, the closest non-relative person NP among all NPs preceding ana from the same sentence as well as from the previous two sentences is chosen as its antecedent.

Rule3: GPE job title NPs. In news articles, it is common that a globally salient geo-political entity (hence GPE, e.g. *Japan* or *U.S.*) is introduced in the beginning, then later a related job title NP (e.g. *officials* or *the prime minister*) is used directly without referring to this GPE explicitly. To resolve such bridging cases accurately, we compile a list of twelve job titles which are related to GPEs (e.g. *mayor* or *official*). An NP from A is added to A_{r3} if its head appears in this list and does not have a country pre-modification (e.g. *the Egyptian president*). Then for each potential bridging anaphor $ana \in A_{r3}$, the most salient GPE NP among all NPs preceding ana is predicted as its antecedent. We use the NP's frequency in the whole document to measure its salience throughout the paper. In case of a tie, the closest one is chosen to be the predicted antecedent.

Rule4: role NPs. Compared to Rule3, Rule4 is designed to resolve more general role NPs to their implicit possessor arguments. We extract a list containing around 100 nouns which specify professional roles from WordNet (e.g. *chairman*, *president* or *professor*). An NP from A is added to

A_{r4} if its head appears in this list. Then for each potential bridging anaphor $ana \in A_{r4}$, the most salient proper name NP which stands for an organization among all NPs preceding ana from the same sentence as well as from the previous four sentences is chosen as its antecedent (if such an NP exists). Recency is again used to break ties.

Rule5: percentage NPs. In set bridging as shown in Example 7, the anaphor (**Seventeen percent**) is indicated by a percentage expression from A, which is often in the subject position. The antecedent (*the firms*) is predicted to be the closest NP which modifies another percentage NP via the preposition "of" among all NPs occurring in the same or up to two sentences prior to the potential anaphor.

(7) 22% of *the firms* said employees or owners had been robbed on their way to or from work. **Seventeen percent** reported their customers being robbed.

Rule6: other set member NPs. In set bridging, apart from percentage expressions, numbers or indefinite pronouns are also good indicators for bridging anaphora. For such cases, the anaphor is predicted if it is: (1) a number expression (e.g. **One** in Example 3) or an indefinite pronoun(e.g. **some**, as shown in Example 8) from A; and (2) a subject NP. The antecedent is predicted to be the closest NP among all plural, subject NPs preceding the potential anaphor from the same sentence as well as from the previous two sentences (e.g. *Reds and yellows* in Example 8). If such an NP does not exist, the closest NP among all plural, object NPs preceding the potential anaphor from the same sentence as well as from the previous two sentences is chosen to be the predicted antecedent (e.g. *several problems* in Example 3).

(8) *Reds and yellows* went about their business with a kind of measured grimness. **Some** frantically dumped belongings into pillow-cases.

Rule7: argument-taking NPs I. Laparra and Rigau (2013) found that different instances of the same predicate in a document likely maintain the same argument fillers. Here we follow this assumption but apply it to nouns and their nominal modifiers only: different instances of the same noun predicate likely maintain the same argument fillers indicated by nominal modifiers. First, a common noun phrase from A is added to A_{r7} if: (1) its argument taking ratio α is bigger than 0.5; (2) it does not contain any nominal or adjective pre-modifications; and (3) it is not modified by indefinite determiners[5] which usually introduce new discourse referents (Hawkins, 1978). Then for each potential bridging anaphor $ana \in A_{r7}$, we choose the antecedent by performing the following steps:

1. We take ana's head lemma form ana_h and collect all its syntactic modifications in the document. We consider nominal pre-modification, possessive modification as well as prepositional post-modification. All realizations of these modifications which precede ana form the antecedent candidates set C_{ana}.

2. We choose the most recent NP from C_{ana} as the predicted antecedent for the potential bridging anaphor ana.

In Example 9, we first predict the two occurrences of **residents** as bridging anaphors. Since in the text, other occurrences of the lemma "resident" are modified by "Marina" (supported by Marina residents) and "buildings" (supported by some residents of badly damaged buildings), we collect all NPs whose syntactic head is "Marina" or "buildings" in C_{ana} (i.e. *Marina*, *badly damaged buildings* and *buildings with substantial damage*). Then among all NPs in C_{ana}, the most recent NP is chosen to be the antecedent (i.e. *buildings with substantial damage*).

(9) She finds the response of Marina residents to the devastation of their homes "incredible".
. . .
Out on the streets, some residents of badly damaged buildings were allowed a 15 minute scavenger hunt through their possessions.
. . .
After being inspected, *buildings with substantial damage* were color - coded.
Green allowed **residents** to re-enter; red allowed **residents** one last entry to gather everything they could within 15 minutes.

Rule8: argument-taking NPs II. Prince (1992) found that discourse-old entities are more likely

---

[5] We compile a list of 17 such determiners, such as *a*, *an* or *one*.

to be represented by NPs in subject position. Although she could not draw a similar conclusion when collapsing *Inferrable* (= *bridging*) with *Discourse-old Nonpronominal*, we find that in the development set, an argument-taking NP in the subject position is a good indicator for bridging anaphora (e.g. **participants** in Example 10). A common noun phrase from A is collected in A_{r8} if: (1) its argument taking ratio α is bigger than 0.5; (2) it does not contain any nominal or adjective pre-modifications; and (3) it is in the subject position. Semantic connectivity again is used as the criteria to choose the antecedent: for each potential bridging anaphor $ana \in A_{r8}$, the NP with the strongest semantic connectivity to ana among all NPs preceding ana from the same sentence as well as from the previous two sentences is predicted to be the antecedent.

(10) Initial steps were taken at *Poland's first international environmental conference, which I attended last month.* ... While Polish data have been freely available since 1980, it was no accident that **participants** urged the free flow of information.

3.2 Post-processing

In the bridging link prediction component, each rule is applied separately. To resolve the conflicts between different rules (e.g., two rules predict different antecedents for the same potential anaphor), a post processing step is applied. We first order the rules according to their precision for predicting bridging pairs (i.e., recognizing bridging anaphors and finding links to antecedents) in the development set. When a conflict happens, the rule with the highest order has the priority to decide the antecedent. Table 2 summarizes the rules described in Section 3.1, the numbers in square brackets in the first column indicate the order of the rules. Table 3 shows the precisions of bridging anaphora recognition and bridging pairs prediction for each rule in the development set. Firing rate is the proportion of bridging links predicted by rule r among all predicted links.

4 Experiments and Results

4.1 Experimental Setup

We conduct all experiments on the ISNotes corpus. We use the OntoNotes named entity and syntactic annotations to extract features. Ten documents containing 113 bridging anaphors from the ISNotes corpus are set as the development set to estimate parameters for the rule-based system. The remaining 40 documents are used as the test set. In order to compare the results of different systems directly, we evaluate all systems on the test set.

4.2 Evaluation Metric

In ISNotes, bridging is annotated mostly between an NP (anaphor) and an entity (antecedent)[6], so that a bridging anaphor could have multiple links to different instantiations of the same entity (entity information is based on the Ontonotes coreference annotation). For bridging resolution, we use an evaluation metric based on bridging anaphors instead of all links between bridging anaphors and their antecedent instantiations. A link predicted by the system is counted as correct if it recognizes the bridging anaphor correctly and links the anaphor to any instantiation of the right antecedent entity preceding the anaphor.

In the evaluation metric, recall is calculated via the number of the correct links predicted by the system (one unique link per each predicted anaphor) divided by the total number of the gold bridging anaphors, precision is calculated via the number of the correct links predicted by the system divided by the total links predicted by the system.

4.3 A Learning-based Approach

To compare our rule-based system (hence *ruleSystem*, described in Section 3) with other approaches, we implement a learning-based system for unrestricted bridging resolution. We adapt the pairwise model which is widely used in coreference resolution (Soon et al., 2001). Similar to the rule-based system, we first create an initial list of possible bridging anaphora A_{ml} with one more constraint. The purpose is to exclude as many obvious non-bridging anaphoric NPs from the list as possible. An NP is added to A_{ml} if: (1) it does not contain any other NPs; (2) it is not modified by pre-modifications which strongly indicate comparative NPs; and (3) it is not a pronoun or a proper name. Then for each NP $a \in A_{ml}$, a list of antecedent candidates C_a is created by including all NPs preceding a from the same sentence

---

[6]There are a few cases where bridging is annotated between an NP and a non-NP antecedent (e.g. *verbs* or *clauses*).

rule	anaphor	antecedent	antecedent candidates scope
rule1 [2]	building part NPs	the NP with the strongest semantic connectivity to the potential anaphor	two
rule2 [5]	relative person NPs	the closest person NP which is not a relative NP	two
rule3 [6]	GPE job title NPs	the most salient GPE NP	all
rule4 [7]	role NPs	the most salient organization NP	four
rule5 [1]	percentage NPs	the closest NP which modifies another percentage NP via the preposition "of"	two
rule6 [3]	other set member NPs	the closest subject, plural NP; otherwise the closest object, plural NP	two
rule7 [4]	argument-taking NPs I	the closest NP whose head is an unfilled role of the potential anaphor (such a role is predicted via syntactic modifications of NPs which have the same head as the potential anaphor)	all
rule8 [8]	argument-taking NPs II	the NP with the strongest semantic connectivity to the potential anaphor	two

Table 2: Rules for unrestricted bridging resolution. Antecedent candidates scope are verified in the development set: "all" represents all NPs preceding the potential anaphor from the whole document, "four" NPs occurring in the same or up to four sentences prior to the potential anaphor, "two" NPs occurring in the same or up to two sentences prior to the potential anaphor.

rule	anaphora	anaphora recognition precision	bridging pairs prediction precision	firing rate
rule1 [2]	building part NPs	75.0%	50.0%	6.1%
rule2 [5]	relative person NPs	69.2%	46.2%	6.1%
rule3 [6]	GPE job title NPs	52.6%	44.7%	19.4%
rule4 [7]	role NPs	61.7%	32.1%	28.6%
rule5 [1]	percentage NPs	100.0%	100.0%	2.6%
rule6 [3]	other set member NPs	66.7%	46.7%	7.8%
rule7 [4]	argument-taking NPs I	53.8%	46.4%	6.1%
rule8 [8]	argument-taking NPs II	64.5%	25.0%	25.5%

Table 3: Precision of bridging anaphora recognition and bridging pairs prediction for each rule in the development set. The numbers in square brackets in the first column indicate the order of the rules.

as well as from the previous two sentences[7]. We create a pairwise instance (a, c) for every $c \in C_a$. We also add extra pairwise instances from the prediction of *ruleSystem* to the learning-based system. In the decoding stage, the *best first* strategy (Ng and Cardie, 2002) is used to predict the bridging links. Specifically, for each $a \in A_{ml}$, we predict the bridging link to be the most confident pair (a, c_{ante}) among all instances with the positive prediction. We use SVM^{light} to conduct the experiments[8]. All experiments are conducted via 10-fold cross-validation on the whole corpus[9].

mlSystem_ruleFeats We provide *mlSystem_ruleFeats* with the same knowledge resources as the rule-based system. All rules from the rule-based system are incorporated into *mlSystem_ruleFeats* as the features.

mlSystem_ruleFeats + atomFeats We augment *mlSystem_ruleFeats* with more features from our previous work (Markert et al., 2012; Hou et al., 2013a; Hou et al., 2013b) on bridging anaphora recognition and antecedent selection. Some of these features overlap with the atomic features used in the rule-based system.

Table 4 shows all the features we use for recognizing bridging anaphora. "*" indicates the resources are used in the rule-based system. We apply them to the first element a of a pairwise instance (a, c). Markert et al. (2012) and Hou et

---

[7]In ISNotes, 71% of NP antecedents occur in the same or up to two sentences prior to the anaphor. Initial experiments show that increasing the window size more than two sentences decreases the performance.

[8]To deal with data imbalance, the SVM^{light} parameter is set according to the ratio between positive and negative instances in the training set.

[9]To compare the learning-based approach to the rule-based system described in Section 3 directly, we report the

results of learning-based approaches on the same test set as the rule-based system.

Markert et al. local feature set

$f1$ FullPrevMention (b) $*$	$f2$ FullPreMentionTime (n)	$f3$ PartialPreMention (b)
$f4$ ContentWordPreMention (b)	$f5$ Determiner (n) $*$	$f6$ NPtype (n) $*$
$f7$ NPlength (int)	$f8$ GrammaticalRole (n) $*$	$f9$ NPNumber (n) $*$
$f10$ PreModByCompMarker (b) $*$		

Hou et al. local feature set

features to identify bridging anaphora

$f1$ IsCoherenceGap (b)	$f2$ IsSentFirstMention (b)	$f3$ IsDocFirstMention (b)
$f4$ IsWordNetRelationalNoun (b) $*$	$f5$ IsInquirerRoleNoun (b)	$f6$ IsBuildingPart (b) $*$
$f7$ IsSetElement (b) $*$	$f8$ PreModSpatialTemporal (b)	$f9$ IsYearExpression (b)
$f10$ PreModifiedByCountry (b) $*$	$f11$ AppearInIfClause (b)	$f12$ VerbPosTag (l)
$f13$ IsFrequentGenericNP (b)	$f14$ WorldKnowledgeNP (l)	$f15$ Unigrams (l)
$f16$ PreModByGeneralQuantifier (b)	$f17$ BridgingHeadNP (l)	$f18$ HasChildNP (b) $*$

features to identify function and worldKnowledge NPs

$f20$ DependOnChangeVerb (b)	$f21$ IsFrequentProperName (b)

Table 4: Features for bridging anaphora recognition from Markert et al. (2012) and Hou et al. (2013a). "b" indicates binary, "n" nominal, "l" lexical features, "$*$" resources used in the rule-based system.

Group	Feature	Value
semantic	$f1$ preposition pattern $*$	the normalized hit counts of the preposition pattern query *a prep. c* (e.g. *participants of the conference*) in big corpora
	$f2$ verb pattern	the normalized hit counts of the verb pattern query *c verb$_a$* or *verb$_a$ c* in big corpora (for set bridging in Example 7, the pattern query is *the firms reported*)
	$f3$ WordNet partOf	whether a partOf relation holds between a and c in WordNet
	$f4$ semantic class $*$	16 classes, e.g. location, organization, GPE, rolePerson, relativePerson, product, date, money, percent
salience	$f5$ document span	the normalized value of the span of text in which c is mentioned
	$f6$ utterance distance	the sentence distance between a and c
	$f7$ local first mention	whether c is the first mention within the previous five sentences
	$f8$ global first mention	whether c is the first mention in the whole document
syntactic & lexical	$f9$ isSameHead	whether a and c share the same head (exclude coreferent antecedent candidates)
	$f10$ isWordOverlap	whether a is prenominally modified by the head of c (for bridging where the anaphor is a compound noun, such as *the mine-**mine security***)
	$f11$ isCoArgument	whether subject c and object a are dependent on the same verb (the subject can not be the bridging antecedent of the object in the same clause)
	$f12$ WordNet distance	the inverse value of the shortest path length between a and c in WordNet

Table 5: Features for antecedent selection from Hou et al. (2013b). "$*$" indicates resources used in the rule-based system.

al. (2013a) classify eight fine-grained information status (IS) categories for NPs: *old*, *new* and 6 *mediated* categories (*syntactic*, *worldKnowledge*, *bridging*, *comparative*, *aggregate* and *function*). Features from Markert et al. (2012) work well to identify *old*, *new* and several *mediated* categories but fail to recognize most bridging anaphora. Hou et al. (2013a) remedy this by adding discourse structure features ($f1$-$f3$), semantic features ($f4$-$f10$) and features to detect generic nouns ($f11$-

Feature	Value
for anaphor candidate a	
$f1$ preModByNominal	whether a contains any nominal pre-modifications
$f2$ preModByAdj	whether a contains any adjective modifications
$f3$ isGPEJobTitle	whether a is a job title about GPE (e.g. *mayor* or *official*)
$f4$ isArgumentTakingNP	whether the argument taking ratio of a is bigger than 0.5
for antecedent candidate c	
$f5$ fullMentionTime	the normalized value of the frequency of c in the whole document
for pairwise instance (a, c)	
$f6$ word distance	the token distance between a and c

Table 6: Additional atomic features from the rule-based system.

$f14$ and $f16$).

Table 5 shows all features we use for selecting antecedents for bridging anaphora. "∗" indicates the resources that are used in the rule-based system. These features are from Hou et al. (2013b)'s local pairwise model. They try to model: (1) the semantic relations between bridging anaphors and their antecedents ($f1$ to $f4$); (2) the salience of an antecedent from different perspectives ($f5$ to $f8$); and (3) the syntactic and lexical constraints between anaphor and antecedent ($f9$ to $f12$).

Apart from the features shown in Table 4 and Table 5, we further enrich *mlSystem_ruleFeats+atomFeats* with additional atomic features used in the rule-based system (Table 6).

mlSystem_atomFeats Based on *mlSystem_ruleFeats+atomFeats*, the rule features from the rule-based system are removed.

4.4 Baseline

We also reimplement the rule-based system from Vieira and Poesio (2000) as a baseline. The original algorithm focuses on processing definite NPs. It classifies four categories for the definite NPs: *discourse new*, *direct anaphora* (same-head coreferent anaphora), *lenient bridging* and *Unknown*. This algorithm also finds antecedents for NPs which belong to *direct anaphora* or *lenient bridging*.

Since Vieira and Poesio (2000) include *different-head coreference* into their *lenient bridging* category, we further divide their *lenient bridging* category into two subcategories: *different-head coreference* and *bridging*. Figure 1 shows the details of the division after failing to classify an NP as *discourse new* or *direct anaphora*. For more details about the whole system, see Vieira and Poesio (2000). We then apply this slightly revised algorithm to process all NPs in the initial list of potential bridging anaphora A from *ruleSystem* (described in Section 3.1).

4.5 Results and Discussion

Table 7 shows the results on the same test set of different approaches for unrestricted bridging resolution. The results reveal the difficulty of the task, when evaluating on a realistic scenario without constraints on types of bridging anaphora and bridging relations.

Both our rule-based system and all learning-based approaches significantly outperform the baseline at $p < 0.01$ (randomization test). The low recall in *baseline* is predictable, since it only considers meronymy bridging and compound noun anaphors whose head is prenominally modified by the antecedent head. (e.g. *the state–* **state gasoline taxes**). Under the same features, the learning-based approach (*mlSystem_ruleFeats*) performs slightly worse than the rule-based system (*ruleSystem*) with regard to the F-score.

	R	P	F
baseline	2.9	13.3	4.8
ruleSystem	11.9	42.9	**18.6**
mlSystem_ruleFeats	12.1	35.0	18.0
mlSystem_ruleFeats+atomFeats	16.7	21.2	**18.7**
mlSystem_atomFeats	20.5	10.1	13.5

Table 7: Experimental results for the baseline, the rule-based system and the learning-based systems.

Surprisingly, incorporating rich features into the learning-based approach (*mlSystem_ruleFeats+atomFeats*) does not yield much improvement over the rule-based system (with an
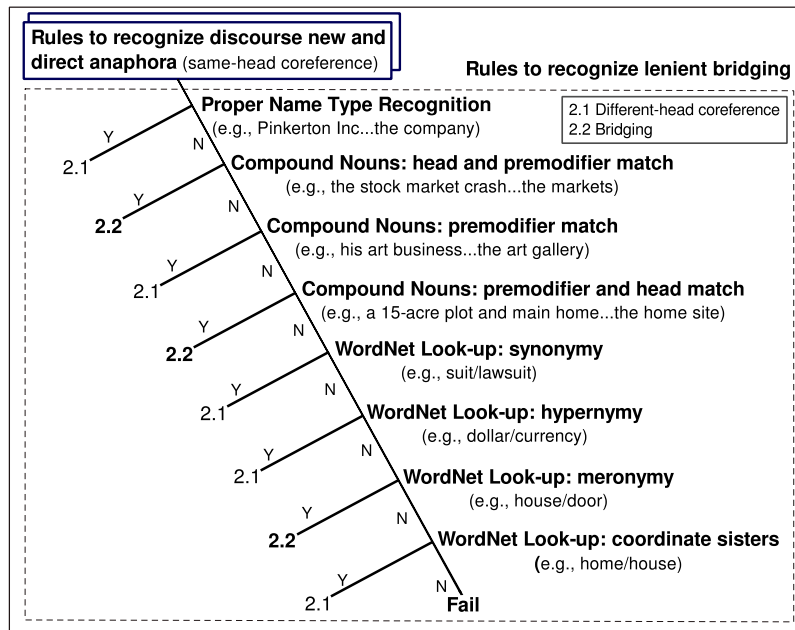
Figure 1: Vieria & Poesio's (2000) original algorithm for processing definite NPs. We further divide their lenient bridging category into two subcategories: *2.1 Different-head coreference* and *2.2 Bridging*.

F-score of 18.7 in *mlSystem_ruleFeats+atomFeats* compared to 18.6 in *ruleSystem*). We suppose that the learning-based system generalizes poorly with only atomic features in Table 4, Table 5 and Table 6. Results on *mlSystem_atomFeats* support our assumption: the F-score drops considerably after removing the rule features. Although ISNotes is a reasonably sized corpus for bridging compared to previous work, diverse bridging relations, especially lots of context specific relations such as *pachinko*-**devotees** or *palms*-**the thieves**, lead to relatively small-scale training data for each type of relation. Therefore it is difficult for the learning-based approach to learn effective rules to predict bridging links.

However, all learning-based systems tend to have higher recall but lower precision compared to the rule-based system. This suggests that the learning-based systems are "greedy" to predict bridging links. A close look at these links in *mlSystem_atomFeats* indicates that the learning-based system predicts more correct bridging anaphors but fails to find the correct antecedents. In fact, lots of those "half" correct links sound reasonable without the specific context, such as *the story-readers* (gold bridging link: *this novel*-**readers**) or *the executive director's office-the desks* (gold bridging link: *the fund's building*-**the desks**).

5 Conclusions

We proposed a rule-based approach for unrestricted bridging resolution where bridging anaphora are not limited to definite NPs and the relations between anaphor and antecedent are not restricted to meronymic relations. We designed eight rules to resolve bridging based on linguistic intuition. Our rule-based system performs better than a learning-based approach which has access to the same knowledge resources as the rule-based system. Particularly, the learning-based system enriched with more features does not yield much improvement over the rule-based system. We speculate that the learning-based system could benefit from more training data. Furthermore, better methods to model the semantics of the specific context need to be explored in the future.

This work is – to our knowledge – the first bridging resolution system that handles the unrestricted phenomenon in a realistic setting.

Acknowledgements

References

Nicholas Asher and Alex Lascarides. 1998. Bridging. *Journal of Semantics*, 15:83–113.

Razvan Bunescu. 2003. Associative anaphora resolution: A Web-based approach. In *Proceedings of the EACL 2003 Workshop on The Computational Treatment of Anaphora,* Budapest, Hungary, 14 April, 2003, pages 47–52.

Aoife Cahill and Arndt Riester. 2012. Automatically acquiring fine-grained information status distinctions in German. In *Proceedings of the SIGdial 2012 Conference: The 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue,* Seoul, Korea, 5–6 July 2012, pages 232–236.

Philipp Cimiano. 2006. Ingredients of a first-order account of bridging. In *Proceedings of the 5th International Workshop on Inference in Computational Semantics,* Buxton, U.K., 20–21 April 2006, pages 139–144.

Herbert H. Clark. 1975. Bridging. In *Proceedings of the Conference on Theoretical Issues in Natural Language Processing,* Cambridge, Mass., June 1975, pages 169–174.

Jeanette K. Gundel, Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69:274–307.

John A. Hawkins. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. Humanities Press, Atlantic Highlands, N.J.

Yufang Hou, Katja Markert, and Michael Strube. 2013a. Cascading collective classification for bridging anaphora recognition using a rich linguistic feature set. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing,* Seattle, Wash., 18–21 October 2013, pages 814–820.

Yufang Hou, Katja Markert, and Michael Strube. 2013b. Global inference for bridging anaphora resolution. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* Atlanta, Georgia, 9–14 June 2013, pages 907–917.

Egoitz Laparra and German Rigau. 2013. ImpAr: A deterministic algorithm for implicit semantic role labelling. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics,* Sofia, Bulgaria, 4–9 August 2013, pages 1180–1189.

Emmanuel Lassalle and Pascal Denis. 2011. Leveraging different meronym discovery methods for bridging resolution in French. In *Proceedings of the 8th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2011),* Faro, Algarve, Portugal, 6–7 October 2011, pages 35–46.

Sebastian Löbner. 1985. Definites. *Journal of Semantics*, 4:279–326.

Sebastian Löbner. 1998. Definite associative anaphora. Unpublished Manuscript, Heinrich-Heine-Universität Düsseldorf.

Katja Markert, Malvina Nissim, and Natalia N. Modjeska. 2003. Using the web for nominal anaphora resolution. In *Proceedings of the EACL Workshop on the Computational Treatment of Anaphora.* Budapest, Hungary, 14 April 2003, pages 39–46.

Katja Markert, Yufang Hou, and Michael Strube. 2012. Collective classification for fine-grained information status. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics,* Jeju Island, Korea, 8–14 July 2012, pages 795–804.

Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely, Veronika Zielinska, Brian Young, and Ralph Grishaman. 2004. Annotating noun argument structure for NomBank. In *Proceedings of the 4th International Conference on Language Resources and Evaluation,* Lisbon, Portugal, 26–28 May 2004, pages 803–806.

Natalia M. Modjeska, Katja Markert, and Malvina Nissim. 2003. Using the web in machine learning for other-anaphora resolution. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing,* Sapporo, Japan, 11–12 July 2003, pages 176–183.

Vincent Ng and Claire Cardie. 2002. Improving machine learning approaches to coreference resolution. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics,* Philadelphia, Penn., 7–12 July 2002, pages 104–111.

Robert Parker, David Graff, Junbo Kong, Ke Chen, and Kazuaki Maeda. 2011. English Gigaword Fifth Edition. LDC2011T07.

Massimo Poesio and Renata Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2):183–216.

Massimo Poesio, Renata Vieira, and Simone Teufel. 1997. Resolving bridging references in unrestricted text. In *Proceedings of the ACL Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Text, Madrid, Spain, July 1997*, pages 1–6.

Massimo Poesio, Rahul Mehta, Axel Maroudas, and Janet Hitzeman. 2004. Learning to resolve bridging references. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics,* Barcelona, Spain, 21–26 July 2004, pages 143–150.

Ellen F. Prince. 1981. Towards a taxonomy of given-new information. In P. Cole, editor, *Radical Pragmatics*, pages 223–255. Academic Press, New York, N.Y.

Ellen F. Prince. 1992. The ZPG letter: Subjects, definiteness, and information-status. In W.C. Mann and S.A. Thompson, editors, *Discourse Description. Diverse Linguistic Analyses of a Fund-Raising Text*, pages 295–325. John Benjamins, Amsterdam.

Altaf Rahman and Vincent Ng. 2012. Learning the fine-grained information status of discourse entities. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics,* Avignon, France, 23–27 April 2012, pages 798–807.

Ina Rösiger and Simone Teufel. 2014. Resolving coreference and associative noun phrases in scientific text. In *Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics,* Gothenburg, Sweden, 26–30 April 2014, pages 44–55.

Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yong Lim. 2001. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, 27(4):521–544.

Philip J. Stone, Dexter C. Dunphy, Marshall S. Smith, Daniel M. Ogilvie, and Cambridge Computer Associates. 1966. *General Inquirer: A Computer Approach to Content Analysis*. MIT Press, Cambridge, Mass.

Renata Vieira and Massimo Poesio. 2000. An empirically-based system for processing definite descriptions. *Computational Linguistics*, 26(4):539–593.

Ralph Weischedel, Martha Palmer, Mitchell Marcus, Eduard Hovy, Sameer Pradhan, Lance Ramshaw, Nianwen Xue, Ann Taylor, Jeff Kaufman, Michelle Franchini, Mohammed El-Bachouti, Robert Belvin, and Ann Houston. 2011. OntoNotes release 4.0. LDC2011T03, Philadelphia, Penn.: Linguistic Data Consortium.