

Leveraging Verb-Argument Structures to Infer Semantic Relations

Eduardo Blanco and Dan Moldovan

Lymba Corporation

Richardson, TX 75080 USA

{eduardo,moldovan}@lymba.com

Abstract

This paper presents a methodology to infer implicit semantic relations from verb-argument structures. An annotation effort shows implicit relations boost the amount of meaning explicitly encoded for verbs. Experimental results with automatically obtained parse trees and verb-argument structures demonstrate that inferring implicit relations is a doable task.

1 Introduction

Automatic extraction of semantic relations is an important step towards capturing the meaning of text. Semantic relations explicitly encode links between concepts. For example, in *The accident left him a changed man*, the ‘accident’ is the CAUSE of the man undergoing some ‘change’. A question answering system would benefit from detecting this relation when answering *Why did he change?*

Extracting *all* semantic relations from text is a monumental task and is at the core of language understanding. In recent years, approaches that aim at extracting a subset of all relations have achieved great success. In particular, previous research (Carreras and Màrquez, 2005; Punyakanok et al., 2008; Che et al., 2010; Zapirain et al., 2010) focused on verb-argument structures, i.e., relations between a verb and its syntactic arguments. PropBank (Palmer et al., 2005) is the corpus of reference for verb-argument relations. However, relations between a verb and its syntactic arguments are only a fraction of the relations present in texts.

Consider the statement $[Mr. Brown]_{NP_1}$ *succeeds* $[Joseph W. Hibben, who retired last August]_{NP_2}$ and its parse tree (Figure 1). Verb-argument relations encode that NP_1 is the AGENT and NP_2 is the THEME of verb ‘succeeds’ (PropBank uses labels ARG_0 and ARG_1). Any semantic relation between ‘succeeds’ and concepts dominated in the parse tree by one of its syntactic arguments NP_1 or NP_2 , e.g., ‘succeeds’ oc-

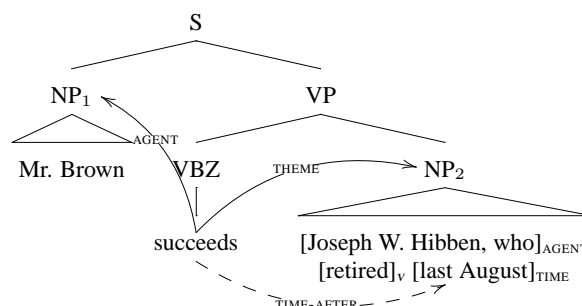


Figure 1: Example of parse tree and verb-argument structures (solid arrows). The relation between ‘succeeds’ and ‘last August’ is missing, but a TIME-AFTER holds (dashed arrow).

curred after ‘last August’, are missing. Note that in this example, verb-argument structures encode that ‘retired’ has TIME ‘last August’, and this knowledge could be exploited to infer the missing relation. The work presented here stems from two observations: (1) verbs are semantically connected with concepts that are not direct syntactic arguments (henceforth, *implicit relations*); and (2) verb-argument structures can be leveraged to infer implicit relations.

This paper goes beyond verb-argument structures and targets implicit relations like the one depicted above. TIME, LOCATION, MANNER, PURPOSE and CAUSE are inferred without imposing syntactic restrictions between their arguments: systems trained over PropBank do not attempt to extract these relations. An annotation effort demonstrates implicit relations reveal as much as 30% of meaning on top of verb-argument structures. The main contributions are: (1) empirical study of verb-argument structures and implicit relations in PropBank; (2) annotations of implicit relations on top of PropBank; (3) novel features extracted from verb-argument structures; and (4) experimental results with features derived from gold and automatically obtained linguistic information, showing implicit relations can be extracted in a realistic environment.

2 Related Work

Several systems to extract verb-argument structures from plain text have been proposed (Johansson and Nugues, 2008; Che et al., 2010). The work presented here complements them with additional semantic relations. The TimeBank corpus (Pustejovsky et al., 2003) and TempEval competitions (UzZaman et al., 2013) target events and detailed temporal information; this work also targets LOCATION, MANNER, PURPOSE and CAUSE. Extracting missing relations is not a new problem. Early work focused on a very limited domain (Palmer et al., 1986; Tetreault, 2002) or did not attempt to automate the task (Whittemore et al., 1991). This section focuses on more recent work.

Gerber and Chai (2010) augment NomBank annotations (Meyers et al., 2004) of 10 predicates with additional core arguments. Their supervised systems obtain F-measures of 42.3 and 50.3 (Gerber and Chai, 2012). Laparra and Rigau (2013a) present a deterministic algorithm and obtain an F-measure of 45.3. In contrast, our approach does not focus on a few selected predicates or core arguments. It targets *all predicates* and argument modifiers (AM-TMP, AM-MNR, AM-LOC, etc.), whose meaning is shared across verbs.

The SemEval-2010 Task 10: Linking Events and their Participants in Discourse (Ruppenhofer et al., 2009) targeted cross-sentence missing core arguments in both PropBank and FrameNet (Baker et al., 1998). Ruppenhofer et al. (2013) detail the annotations and results. The task proved extremely difficult, participants (Chen et al., 2010; Tonelli and Delmonte, 2010) reported overall F-measures around 2 (out of 100). Posterior work (Silberer and Frank, 2012; Laparra and Rigau, 2013b) reported F-measures below 20 for the same task. The work presented here does not target missing core arguments but modifiers within the same sentence. Furthermore, results show our proposal is useful in a real environment.

Finally, our previous work (Blanco and Moldovan, 2011; Blanco and Moldovan, 2014) proposed composing new relations out of chains of previously extracted relations. This approach is unsupervised and accurate (88% with gold annotations), but inferences are made only between the ends of chains of existing relations. Our current proposal also leverages relations previously extracted, but productivity is higher and results with automatic annotations are presented.

[But] _{MDis} [the surprisingly durable seven-year economic expansion] _{ARG0} has [made] _v [mincemeat] _{ARG1} [of more than one forecast] _{ARG2} .
Also, financial planners advising on insurance say that to their knowledge there has not yet been [a tax ruling] _{ARG0} [exempting] _v [these advance payments] _{ARG1} [from taxes] _{ARG2} .

Table 1: Examples of verb-argument structures from PropBank.

3 Verb-Argument Structures and Implicit Relations

Throughout this paper, $R(x, y)$ denotes a semantic relation R holding between x and y . $R(x, y)$ is interpreted “ x has R y ”, e.g., $AGENT(took, Bill)$ could be read “*took* has $AGENT$ *Bill*”. Verb-argument structures, or semantic roles, account for semantic relations between a verb and its syntactic arguments. In other words, $R(x, y)$ is a semantic role if ‘ x ’ is a verb and ‘ y ’ a syntactic argument of ‘ x ’, and all semantic roles with ‘ x ’ as first argument form the verb-argument structure of verb ‘ x ’. Implicit relations are relations $R(x, y)$ where x is a verb and y is not a syntactic argument of x .

The work presented in this paper aims at complementing verb-argument structures with implicit semantic relations. We follow a practical approach by inferring implicit relations from PropBank’s verb-argument structures. We believe this is an advantage since PropBank is well-known in the field and several tools to predict PropBank annotations are documented and publicly available.¹ The work presented here could be incorporated in any NLP pipeline after role labeling without modifications to other components. Furthermore, working on top of PropBank allows us to quantify the impact of features derived from gold and automatically extracted linguistic information when inferring implicit relations (Section 6).

3.1 Verb-Argument structures in PropBank

PropBank (Palmer et al., 2005) annotates verb-argument structures on top of the syntactic trees of the Penn TreeBank (Marcus et al., 1994). It uses a set of numbered arguments² (ARG_0 , ARG_1 , ARG_2 , etc.) and modifiers (AM-TMP, AM-MNR, etc.). Numbered arguments do not share a common meaning across verbs, they are defined on a

¹E.g., Illinois SRL, <http://cogcomp.cs.illinois.edu/page/software>; SENNA, <http://ml.nec-labs.com/senna/>; SwiRL, <http://www.surdeanu.info/mihai/swirl/>

²Numbered arguments are also referred to as *core*.

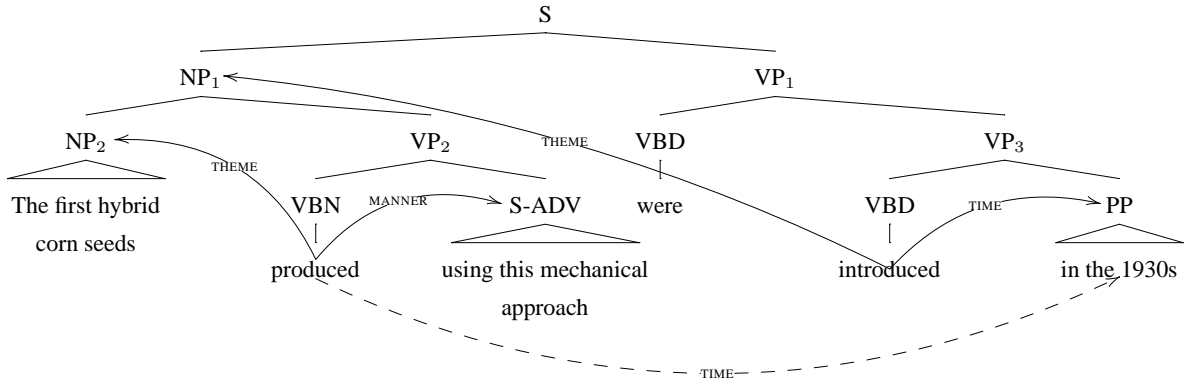


Figure 2: Verb-argument structures (solid arrows) and inferred implicit semantic relation (dashed arrow).

AM-LOC: location	AM-CAU: cause
AM-EXT: extent	AM-TMP: time
AM-DIS: discourse connective	AM-PNC: purpose
AM-ADV: general-purpose	AM-MNR: manner
AM-NEG: negation marker	AM-DIR: direction
AM-MOD: modal verb	

Table 2: Argument modifiers in PropBank.

Label	# predicates	% predicates
ARG ₀	79,334	70.26%
ARG ₁	106,331	94.17%
ARG ₂	24,560	21.75%
AM-TMP	19,756	17.50%
AM-MNR	7,833	6.94%
AM-LOC	7,198	6.37%
AM-PNC	2,784	2.47%
AM-CAU	1,563	1.38%

Table 3: Counts of selected PropBank semantic roles. Total number of predicates is 112,917.

verb by verb basis in each frameset. For example, ARG₂ is used to indicate “*created-from, thing changed*” with verb *make* and “*entity exempted from*” with verb *exempt* (Table 1).

Unlike numbered arguments, modifiers share a common meaning across verbs (Table 2). Some modifiers are arguably not a semantic relation and are not present in most relation inventories (Tratz and Hovy, 2010; Hendrickx et al., 2009). For example, AM-NEG and AM-MOD signal the presence of negation and modals, e.g., $[wo]_{AM-MOD}[n't]_{AM-NEG}[go]_V$. For more information about PropBank annotations and examples, refer to the annotation guidelines.³

Inspecting PropBank annotations one can easily conclude that numbered arguments dominate the annotations and only a few modifiers are an-

³<http://verbs.colorado.edu/~mpalmer/projects/ace/PBguidelines.pdf>

notated (Table 3). ARG₀ and ARG₁ are present in most verb-argument structures, other numbered arguments are often not defined in the corresponding frameset and are thus not annotated.

Examining PropBank one can also conclude that information regarding TIME, LOCATION, MANNER, CAUSE and PURPOSE for a given verb is often present, yet not annotated because the text encoding this knowledge is not a direct syntactic argument of the verb (Section 4.3). Because of this fact, we decided to focus on these five relations.

3.2 Implicit relations in PropBank

Two scenarios are possible when inferring an implicit relation $R(x, y)$: (1) a semantic role $R'(x', y)$ exists; or (2) such a semantic role does not exist. In (1), y is a syntactic argument of some verb x' , where $x \neq x'$ and in (2) that is not the case. Inferences under scenario (1) can be further classified into (1a) when a semantic role $R''(x, y')$ such that y' contains y exists; or (1b) when such a semantic role does not exist. The remainder of this section exemplifies the three scenarios.

The example in Figure 1 falls under scenario (1a). Semantic roles encode, among others, ‘*retired*’ has TIME ‘*last August*’, and ‘*succeeds*’ has AGENT ‘*Mr. Brown*’ and THEME ‘*Joseph W. Hibben, who retired last August*’. The second argument of implicit relation TIME-AFTER(*succeeds*, *last August*) is a semantic role of ‘*retired*’ and is contained in the THEME of ‘*succeeds*’.

Figure 2 shows a statement in which implicit relation TIME(*produced*, *in the 1930s*) could be inferred under scenario (1b). Semantic roles of ‘*produced*’ only indicate that NP₂ is the THEME and S-ADV the MANNER; roles of ‘*introduced*’ indicate that NP₁ is the THEME and ‘*[in the 1930s]_{PP}*’ the TIME. In this case, there is no connection be-

```

rs = {TIME, LOCATION, MANNER, CAUSE, PURPOSE};
foreach semantic role  $R'(x', y)$  such that  $R' \in rs$  do
  foreach verb  $x$  in the same sentence do
    generate potential implicit relation  $R(x, y)$ ;

```

Algorithm 1. Procedure to generate all potential implicit relations in scenario (1) (Section 3.2).

tween ‘produced’ and ‘[in the 1930s]PP’ or any other node subsuming this PP in the parse tree.

Scenario (2) occurs whenever the second argument of implicit relation $R(x, y)$ is not a syntactic argument of a verb. If it were, a semantic role $R'(x', y)$ would exist and it would fall under scenario (1). For example, in $[I]_{AGENT}$ [gave]_v [her]_{RECIPIENT} [a book from 1945]_{THEME}, we could infer the implicit semantic relation “gave occurred after 1945”.

4 Annotating Implicit Relations

Inferring all implicit semantic relations is a challenging task. This paper targets implicit relations that can be inferred under scenarios (1a, 1b); scenario (2) is reserved for future work. All potential implicit relations under scenario (1) are generated using Algorithm 1. A manual annotation effort discards potential implicit relations that do not hold in order to create a gold standard.

4.1 Annotation Guidelines

Annotators are faced with the task of deciding whether a potential implicit relation $R(x, y)$ holds. If it does, they mark it with YES, otherwise with NO. Annotators were initially trained with the original PropBank annotation guidelines⁴ as this task is very similar to annotating PropBank semantic roles. Indeed, the only difference is that ‘y’ is not a syntactic argument of ‘x’.

After some preliminary annotations, we found it useful to account for three subtypes of TIME. This way, richer semantic connections are inferred. When the task is to decide whether implicit relation $TIME(x, y)$ holds, annotators have four labels to choose from: (1) TIME-BEFORE: x occurred before y ; (2) TIME-AFTER: x occurred after y ; (3) TIME-SAME x occurred at/during y ; and (4) NO: y does not describe temporal information of x . If more than one label is valid, annotators choose the one encoding the temporal context y of x starting the earliest. Namely, TIME-BEFORE

⁴<http://verbs.colorado.edu/~mpalmer/projects/ace/PBguidelines.pdf>

has the highest priority, followed by TIME-SAME, TIME-AFTER and finally NO.

Annotation examples are detailed in Section 4.2, the more complex annotations involving TIME are illustrated below. Consider the following statement and PropBank annotations:

[The government’s decision]_{ARG₂, v₁}
 [reflects]_{v₁} [their true desires before
 [the next election]_{ARG₁, v₂}, [expected]_{v₂}
 [in late 1991]_{TIME, v₂}]_{ARG₁, v₁}.

When annotating potential implicit semantic relation $R(\textit{reflects}, \textit{in late 1991})$, annotators may select TIME-BEFORE, TIME-SAME and TIME-AFTER. However, they select TIME-BEFORE because it indicates the temporal context of ‘reflects’ that starts the earliest.

4.2 Annotation Examples

Several annotations examples are shown in Table 4. Semantic roles for statement (1) include $TIME(\textit{remain}, \textit{in 1990})$, $MANNER(\textit{remain}, \textit{at about 1,200 cars})$ and no other TIME or MANNER. Implicit relations reveal two extra semantic connections: $TIME-BEFORE(\textit{said}, \textit{in 1990})$ and $TIME-BEFORE(\textit{expects}, \textit{in 1990})$, i.e., ‘said’ and ‘expects’ occurred before ‘1990’. The potential implicit relations $MANNER(\textit{said}, \textit{at about 1,200 cars})$ and $MANNER(\textit{expects}, \textit{at about 1,200 cars})$ do not hold and are annotated N.

Interpreting statement (2) one can see that ‘this past summer’ is not only indicating the TIME of ‘proposed’; events encoded by verbs ‘make’ and ‘exempt’ occurred after ‘this past summer’. In this example, two implicit semantic relations are inferred from a single semantic role.

Statement (3) shows that two potential implicit relations $R(x, y)$ and $R(x', y)$ sharing the second argument ‘y’ may be assigned different labels. Regarding time, semantic roles only include $TIME(\textit{report}, \textit{in December})$. Implicit relations add $TIME-BEFORE(\textit{proposed}, \textit{in December})$ and $TIME-SAME(\textit{allow}, \textit{in December})$.

Two implicit LOCATION relations are inferred in statement (4): ‘discovered’ and ‘preserving’ occurred ‘in the test-tube experiments’. The potential implicit relation $LOCATION(\textit{said}, \textit{in the test-tube experiments})$ is discarded (annotated N). Statement (5) shows two potential implicit MANNER that can be inferred. The ‘program’ was ‘aired’ and ‘seen by 12 million viewers’ in the following manner: ‘With Mr. Vila as host’.

Statement	TMP				LOC		MNR		PRP		CAU	
	B	A	S	N	Y	N	Y	N	Y	N	Y	N
1: Rolls-Royce said it expects [its U.S. sales] _{ARG1} to [remain] _v [steady] _{ARG3} [at about 1,200 cars] _{MANNER} [in 1990] _{TIME} .												
- said, [in 1990] _{TIME}	✓	-	-	-								
- expects, [in 1990] _{TIME}	✓	-	-	-								
- said, [at about 1,200 cars] _{MANNER}							-	✓				
- expects, [at about 1,200 cars] _{MANNER}							-	✓				
2: They make the argument in letters to the agency about [rule changes] _{ARG1} [proposed] _v [this past summer] _{TIME} that, among other things, exempt many middle-management executives from government supervision.												
- make, [this past summer] _{TIME}	-	✓	-	-								
- exempt, [this past summer] _{TIME}	-	✓	-	-								
3: The proposed changes also allow [executives] _{ARG0} to [report] _v [exercises of options] _{ARG1} [in December] _{TIME} .												
- proposed, [in December] _{TIME}	✓	-	-	-								
- allow, [in December] _{TIME}	-	-	✓	-								
4: Two Japanese scientists said they discovered [an antibody that] _{ARG0} , [in laboratory test-tube experiments] _{LOCATION} , [kills] _v [AIDS-infected cells] _{ARG1} [while preserving healthy cells] _{TIME} .												
- said, [in laboratory test-tube experiments] _{LOCATION}					-	✓						
- discovered, [in laboratory test-tube experiments] _{LOCATION}					✓	-						
- preserving, [in laboratory test-tube experiments] _{LOCATION}					✓	-						
5: [With Mr. Vila as host] _{MANNER} , “[This Old House] _{ARG1} ” [became] _v [one of the Public Broadcasting Service’s top 10 programs] _{ARG2} , [airing weekly on about 300 of the network’s stations and seen by an average of 12 million viewers] _{AM-ADV} .												
- airing, [With Mr. Vila as host] _{MANNER}							✓	-				
- seen, [With Mr. Vila as host] _{MANNER}							✓	-				
[6: It] _{ARG0} [raised] _v [financing of 300 billion lire] _{ARG1} [for the purchase this summer by another Agnelli-related group of the food concern Galbani S.p.A.] _{PURPOSE} , [by selling a chunk of its IFI shares to Mediobanca S.p.A.] _{MANNER}												
- selling, [for the purchase this summer by another ...] _{PURPOSE}									✓	-		
7: [Greece and Turkey] _{ARG0} , for example, are suspected of [overstating] _v [their arsenals] _{ARG1} [in hopes that they can emerge from the arms-reduction treaty with large remaining forces to deter each other] _{PURPOSE} .												
- suspected, [in hopes that they can emerge from the ...] _{PURPOSE}									-	✓		
8: ...the rationalization that [given the country’s lack of natural resources] _{CAUSE} , [they] _{ARG0} [must] _{AM-MOD} [work] _v [hard] _{MANNER} [to create value through exports] _{ARG1} and buy food with the surplus.												
- create, [given the country’s lack of natural resources] _{CAUSE}											✓	-
- buy, [given the country’s lack of natural resources] _{CAUSE}											✓	-
9: Its third-quarter earnings were lower than analysts had forecast, and the company said [it] _{ARG0} had [lowered] _v [its projections for earnings growth through the end of 1990] _{ARG1} [because of planned price cuts] _{CAUSE} .												
- forecast, [because of planned price cuts] _{CAUSE}											-	✓
- said, [because of planned price cuts] _{CAUSE}											-	✓

Table 4: Examples of potential implicit relations and their annotations. All of them but the ones annotated with N can be inferred. B stands for BEFORE, A for AFTER, S for SAME, N for NO and Y for YES. PropBank semantic roles from which implicit relations are generated are indicated between brackets.

Statement (6, 7) exemplify potential implicit PURPOSE relations. While the ‘selling’ event in statement (6) has as its purpose ‘the purchase [...]’ (label Y), the ‘suspected’ event in statement (7) is clearly not done so that ‘they (Greece and Turkey) can emerge from the [...]’ (label N).

Finally, statements (8, 9) exemplify potential implicit CAUSE relations. In (8), both ‘create’ and ‘buy’ are done due to the ‘country’s lack of natural resources’. However, in (9), the analysts ‘forecasting’ and the company ‘saying’ do not have as their cause ‘planned price cuts’.

4.3 Annotation Analysis

Table 5 shows counts for all potential implicit relations annotated. All labels except N indicate a valid implicit relation. 94.1% of potential implicit relations generated from a TIME semantic role can

be inferred. Other roles yield less inferences in relative terms, but substantial additional meaning: LOCATION 39.4%, MANNER 16.7%, PURPOSE 29.4%, and CAUSE 30.2%.

Two annotators performed the annotations. A simple script generated all potential implicit relations and prompted for a label: BEFORE, AFTER, SAME or NO if the potential implicit relation was generated from a TIME semantic role; YES or NO otherwise. Annotators are not concerned with argument identification, as arguments of implicit relations are retrieved from the verb-argument structures in PropBank (Algorithm 1). This makes the annotation process easier and faster.

Annotation quality was calculated with two agreement coefficients: observed agreement (raw percentage of equal annotations) and Cohen’s κ (Artstein and Poesio, 2008). The actual num-

	Source	No.	Name	Description
basic	x	1,2	word, POS tag	x 's surface form and part-of-speech tag
		3	voice	whether x is in active or passive voice
	y	4,5	first word, POS tag	first word and part of speech tag in y
		6,7	last word, POS tag	last word and part-of-speech tag in y
		8,9	head, POS tag	head of y and its part-of-speech tag
		10–12	node, left and right sibling	syntactic nodes of y , and its left and right siblings
	x, y	13	subcategory	concatenation of y 's children nodes
		14	direction	whether x occurs before or after y
15		subsumer	common syntactic node between x and y	
16		path	syntactic path between x and y	
pred structures	x- ps	17–31	verb semantic roles	flags indicating presence of semantic roles in x - ps
	y- ps	32,33	verb, POS tag	verb in y - ps and its part-of-speech tag
		34	arg label	semantic role between verb in y - ps and y
		35–49	arg semantic roles	flags indicating presence of semantic roles in y - ps
	x- ps , y- ps	50	overlapping semantic role	role R'' linking x and y' , where y' contains y
		51	overlapping head	head of y' in semantic role detected in feature 50
	52	overlapping direct	whether feature 51 is the verb in y - ps	

Table 6: Complete feature set to determine whether a potential implicit semantic relation $R(x, y)$ should be inferred. Second column indicates the source: first or second argument (x, y), or their respective predicate structures (x - ps, y - ps). Features in bold are novel and specially designed for our task.

Label		# instances	% instances
TIME	B	3,033	38.4%
	A	2,886	36.5%
	S	1,514	19.2%
	N	463	5.9%
	All	7,896	100.0%
LOCATION	Y	3,345	39.4%
	N	5,151	60.6%
	All	8,496	100.0%
MANNER	Y	1,600	16.7%
	N	7,987	83.3%
	All	9,587	100.0%
PURPOSE	Y	821	29.4%
	N	1,971	70.6%
	All	2,792	100.0%
CAUSE	Y	404	30.2%
	N	909	69.2%
	All	1,313	100.0%

Table 5: Number of potential implicit relations (instances) annotated and counts for each label. Total number of instances is 30,084.

bers are: 78.16% (observed) / 0.687 (κ) for TIME, 86.63% / 0.733 for LOCATION, 93.02% / 0.782 for MANNER, 88.60% / 0.734 for PURPOSE, and 90.91% / 0.810 for CAUSE. These agreements are either comparable or superior to similar previous annotation efforts. Girju et al. (2007) reported observed agreements between 47.8% and 86.1% when annotating 7 semantic relations between nominals, and Bethard et al. (2008) observed agreements of 81.2% and 77.8% (Kappa: 0.715 and 0.556) when annotating temporal and causal relations between event pairs.

5 Inferring Implicit Relations

Inferring implicit relations is reduced to (1) generating potential implicit relations (Algorithm 1) and (2) labeling them. The second task determines if potential implicit relations should be discarded or inferred, all labels but N indicate potential implicit relations that should be inferred. We follow a standard supervised machine learning approach where each potential implicit relation is an instance.

Instances were divided into training (70%) and test (30%). The feature set (Section 5.1) and model parameters were tuned using 10-fold stratified cross-validation over the training split, and results (Section 6) are reported using the test split. More features than the ones presented were tried and discarded because they did not improve performance, e.g., syntactic path between verbs in the verb-argument structures of x and y , depth of both structures, number of tokens in y .

5.1 Feature Selection

The full set of features to determine whether a potential implicit relation $R(x, y)$ can be inferred is summarized in Table 6. Features are classified into *basic* and *predicate structures*. The former are commonly used by semantic role labelers. The latter exploit the output of role labelers, i.e., verb-argument structures, and, to our knowledge, are novel. Results show *predicate structures* features improve performance (Section 6.2).

Basic features are derived from lexical and syntactic information. We do not elaborate more on

Feat No.	Value
1,2	succeeds, VBZ
3	active
4,5	last, JJ
6,7	August, NNP
8,9	August, NNP
10–12	NP, VBD, nil
13	JJ-NNP
14	after
15	VP
16	VBZ+VP-NP-SBAR-S-VP-NP
17–31	ARG ₀ and ARG ₁ true, rest false
32,33	retired, VBD
34	AM-TMP
35-49	ARG ₀ and AM-TMP true, rest false
50	ARG ₁
51	Hibben
52	false

Table 7: Feature values when deciding if $R(\textit{succeeds}, \textit{last summer})$ can be inferred from the verb-argument structures in Figure 1.

these features, detailed descriptions and examples are provided by Gildea and Jurafsky (2002).

Features (17–52) are derived from the *predicate structures* of x and y and specially defined to infer implicit semantic relations. Features (17–31, 35–49) are flags indicating the presence of semantic roles in the predicate structures of x and y .

Features (32–34) characterize the semantic role $R'(x', y)$ from which the potential implicit relation was generated. They specify verb x' , its part-of-speech, and label R' . Note that x' is not present in the potential implicit relation $R(x, y)$, but incorporating this information helps determining whether a relation actually holds as well as label R (TIME-BEFORE, TIME-AFTER, TIME-SAME, etc.).

Finally, features 50–52 apply to inferences under scenario (1a) (Section 3.2). Feature (50) indicates the semantic role $R''(x, y')$, if any, such that y' contains y . Feature (51) indicates the head of argument y' found in feature (50). Feature (52) captures whether the head calculated in feature (51) is the verb in the predicate structure of y .

Table 7 exemplifies all features when deciding whether $\textit{TIME-AFTER}(\textit{succeeds}, \textit{last August})$ can be inferred from the verb-argument structures in *Mr. Brown succeeds Joseph W. Hibben, who retired last August* (Figure 1). Table 8 provides an additional example for features 50–52.

6 Experiments and Results

Experiments were carried out using Support Vector Machines with RBF kernel as implemented in

Mr. Corr resigned to pursue other interests, the airline said.	
ARG ₀ (<i>resigned, Mr. Corr</i>)	
AM-PNC(<i>resigned, to pursue other interests</i>)	
ARG ₀ (<i>pursue, Mr. Corr</i>)	
ARG ₁ (<i>pursue, other interests</i>)	
ARG ₀ (<i>said, the airline</i>)	
ARG ₁ (<i>said, Mr. Corr resigned to pursue other interests</i>)	
feature 50, overlapping sem rel	ARG ₁
feature 51, overlapping head	resigned
feature 52, overlapping direct	true

Table 8: PropBank roles and values for features (50–52) when predicting potential implicit relation $R(\textit{said}, \textit{to pursue other interests})$, labeled N.

LIBSVM (Chang and Lin, 2011). Parameters α and γ were tuned by grid search using 10-fold cross validation over training instances.

Results are reported using features extracted from gold and automatic annotations. Gold annotations are taken directly from the Penn TreeBank and PropBank. Automatic annotations are obtained with Polaris (Moldovan and Blanco, 2012), a semantic parser that among others is trained with PropBank. Results using gold (automatic) annotations are obtained with a model trained with gold (automatic) annotations.

6.1 Detailed Results

Table 9 presents per-relation and overall results. In general terms, there is a decrease in performance when using automatic annotations. The difference is most noticeable in recall and it is due to missing semantic roles, which in turn are often due to syntactic parsing errors. This is not surprising as in order for an implicit relation $R(x, y)$ to be generated as potential and fed to the learning algorithm for classification, a semantic role $R'(x', y)$ must be extracted first (Algorithm 1). However, using automatic annotations brings very little decrease in precision. This leads to the conclusion that as long as ‘ y ’ is identified as a semantic role of some verb, even if it is mislabeled, one can still infer the right implicit relations. Since results obtained with automatic parse trees and semantic roles are a realistic estimation of performance, the remainder of the discussion focuses on those. Results with gold annotations are provided for informational purposes.

Overall results for inferring implicit semantic relations are encouraging: precision 0.66, recall 0.58 and F-measure 0.616. Direct comparison with previous work is not possible because the implicit relations we aim at inferring have not been considered before. However, we note the top

		gold						automatic					
		basic			basic + ps			basic			basic + ps		
		P	R	F	P	R	F	P	R	F	P	R	F
TIME	B	.66	.72	.689	.72	.74	*.730	.64	.65	.643	.68	.67	.677
	A	.63	.74	.681	.67	.75	.708	.61	.68	.642	.66	.72	.687
	S	.57	.41	.477	.54	.45	.491	.55	.36	.437	.55	.38	.450
LOCATION	Y	.71	.61	.656	.70	.64	.669	.71	.56	.624	.71	.58	.635
MANNER	Y	.65	.38	.480	.60	.45	.514	.54	.45	.489	.64	.41	.500
PURPOSE	Y	.65	.58	.613	.69	.60	.642	.56	.49	.525	.68	.49	.572
CAUSE	Y	.71	.60	.650	.74	.62	.675	.69	.65	.670	.71	.63	.669
All		.66	.61	.625	.67	.64	*.651	.63	.57	.591	.66	.58	*.616

Table 9: Results obtained with the test split using features extracted from gold and automatic annotations, and using basic and *predicate structures* (ps) features. Statistical significance between F-measures using *basic* and *basic + predicate structures* features is indicated with * (confidence 95%).

performer (Koomen et al., 2005) at CoNLL-2005 Shared Task on role labeling obtained the following F-measures when extracting the same relations between a verb and its syntactic arguments: 0.774 (TIME), 0.6033 (LOCATION), 0.5922 (MANNER), 0.4541 (PURPOSE) and 0.5397 (CAUSE).

The most difficult relations are TIME-SAME and MANNER, F-measures are 0.450 and 0.500 respectively. Even when using gold annotations these two relations are challenging: F-measures are 0.491 for TIME-SAME, an increase of 9.1%, and 0.514 for MANNER, an increase of 2.8%. Results show that other relations can be inferred with F-measures between 0.635 and 0.687, the only exception is PURPOSE with an F-measure of 0.572.

6.2 Feature Ablation

Results in Table 9 suggest that while implicit relations can be inferred using *basic* features, it is beneficial to complement them with the novel features derived from *predicate structures*. This is true for all relations except CAUSE when using automatic annotations with a negligible difference of 0.001. When considering all implicit relations, the difference in performance is $0.616 - 0.591 = 0.025$, an increase of 4.2% that is statistically significant (Z-test, confidence 95%).

The positive impact of features derived from *predicate structures* is most noticeable when inferring PURPOSE, with an increase of 8.9% ($0.572 - 0.525 = 0.047$). TIME-BEFORE and TIME-AFTER also benefit, with increases of 5.3% ($0.677 - 0.643 = 0.034$) and 7.0% ($0.687 - 0.642 = 0.045$) respectively. The improvement *predicate structures* features bring is statistically significant when

taking into account all relations (confidence 95%). However, due to the lower number of instances, differences in performance when considering individual relations is not statistically significant.

7 Conclusions

Verb-argument structures, or semantic roles, comprise semantic relations between a verb and its syntactic arguments. The work presented in this paper leverages verb-argument structures to infer implicit semantic relations. A relation $R(x, y)$ is implicit if x is a verb and y is not a syntactic argument of x . The method could be incorporated into any NLP pipeline after role labeling without modifications to other components.

An analysis of verb-argument structures and implicit relations in PropBank has been presented. Out of all potential implicit relations $R(x, y)$, this paper targets those that can be generated from a semantic role $R'(x', y)$, where $x \neq x'$. A manual annotation effort demonstrates implicit relations yield substantial additional meaning. Most of the time (94.1%) a semantic role $TIME(x', y)$ is present, we can infer temporal information for other verbs within the same sentence. Productivity is lower but substantial with other roles: 39.4% (LOCATION), 30.2% (CAUSE), 29.4% (PURPOSE) and 16.7% (MANNER).

Experimental results show that implicit relations can be inferred using automatically obtained parse trees and verb-argument structures. Standard machine learning is used to decide whether a potential implicit relation should be inferred, and novel features characterizing the verb-argument structures we infer from have been proposed.

References

- Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596, December.
- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley framenet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1*, ACL '98, pages 86–90, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Steven Bethard, William Corvey, Sara Klingsenstein, and James H. Martin. 2008. Building a Corpus of Temporal-Causal Structure. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, pages 908–915, Marrakech, Morocco. European Language Resources Association (ELRA).
- Eduardo Blanco and Dan Moldovan. 2011. Un-supervised learning of semantic relation composition. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1456–1465, Portland, Oregon, USA, June. Association for Computational Linguistics.
- Eduardo Blanco and Dan Moldovan. 2014. Composition of semantic relations: Theoretical framework and case study. *ACM Trans. Speech Lang. Process.*, 10(4):17:1–17:36, January.
- Xavier Carreras and Lluís Màrquez. 2005. Introduction to the CoNLL-2005 shared task: semantic role labeling. In *CONLL '05: Proceedings of the Ninth Conference on Computational Natural Language Learning*, pages 152–164, Morristown, NJ, USA. Association for Computational Linguistics.
- Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Wanxiang Che, Ting Liu, and Yongqiang Li. 2010. Improving Semantic Role Labeling with Word Sense. In *The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 246–249, Los Angeles, California, June. Association for Computational Linguistics.
- Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A. Smith. 2010. Semafor: Frame argument resolution with log-linear models. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 264–267, Uppsala, Sweden, July. Association for Computational Linguistics.
- Matthew Gerber and Joyce Chai. 2010. Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 1583–1592, Uppsala, Sweden, July. Association for Computational Linguistics.
- Matthew Gerber and Joyce Chai. 2012. Semantic role labeling of implicit arguments for nominal predicates. *Computational Linguistics*, 38:755–798, 2012.
- Daniel Gildea and Daniel Jurafsky. 2002. Automatic Labeling Of Semantic Roles. *Computational Linguistics*, 28:245–288.
- Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, and Deniz Yuret. 2007. SemEval-2007 Task 04: Classification of Semantic Relations between Nominals. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 13–18, Prague, Czech Republic, June. Association for Computational Linguistics.
- Iris Hendrickx, Su N. Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid, Sebastian Padó, Marco Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 2009. SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals. In *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-2009)*, pages 94–99, Boulder, Colorado, June. Association for Computational Linguistics.
- Richard Johansson and Pierre Nugues. 2008. Dependency-based semantic role labeling of propbank. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP '08*, pages 69–78, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Peter Koomen, Vasin Punyakanok, Dan Roth, and Wen T. Yih. 2005. Generalized Inference with Multiple Semantic Role Labeling Systems. In *Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*, pages 181–184, Ann Arbor, Michigan, June. Association for Computational Linguistics.
- Egoitz Laparra and German Rigau. 2013a. Impar: A deterministic algorithm for implicit semantic role labelling. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1180–1189, Sofia, Bulgaria, August. Association for Computational Linguistics.
- Egoitz Laparra and German Rigau. 2013b. Sources of evidence for implicit argument resolution. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Long Papers*, pages 155–166, Potsdam, Germany, March. Association for Computational Linguistics.
- Mitchel Marcus, Beatrice Santorini, and Mary A. Marcinkiewicz. 1994. Building a large annotated

- corpus of English: The Penn Treebank. *Computational linguistics*, 19(2):313–330.
- A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman. 2004. Annotating Noun Argument Structure for NomBank. In *Proceedings of LREC-2004*, pages 803–806, Lisbon, Portugal.
- Dan Moldovan and Eduardo Blanco. 2012. Polaris: Lymba’s semantic parser. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, pages 66–72, Istanbul, Turkey, May. European Language Resources Association (ELRA).
- Martha S. Palmer, Deborah A. Dahl, Rebecca J. Schiffman, Lynette Hirschman, Marcia Linebarger, and John Dowding. 1986. Recovering implicit information. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pages 10–19, New York, New York, USA, July. Association for Computational Linguistics.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106.
- Vasin Punyakanok, Dan Roth, and Wen T. Yih. 2008. The importance of syntactic parsing and inference in semantic role labeling. *Computational Linguistics*, 34(2):257–287, June.
- James Pustejovsky, Patrick Hanks, Roser Sauri, Andrew See, Robert Gaizauskas, Andrea Setzer, Dragomir Radev, Beth Sundheim, David Day, Lisa Ferro, et al. 2003. The timebank corpus. *Corpus linguistics*, 2003:40.
- Josef Ruppenhofer, Caroline Sporleder, Roser Morante, Collin Baker, and Martha Palmer. 2009. SemEval-2010 Task 10: Linking Events and Their Participants in Discourse. In *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions (SEW-2009)*, pages 106–111, Boulder, Colorado, June. Association for Computational Linguistics.
- Josef Ruppenhofer, Russell Lee-Goldman, Caroline Sporleder, and Roser Morante. 2013. Beyond sentence-level semantic role labeling: linking argument structures in discourse. *Language Resources and Evaluation*, 47(3):695–721.
- Carina Silberer and Anette Frank. 2012. Casting implicit role linking as an anaphora resolution task. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics*, pages 1–10, Montréal, Canada, 7-8 June. Association for Computational Linguistics.
- Joel R Tetreault. 2002. Implicit role reference. In *International Symposium on Reference Resolution for Natural Language Processing*, pages 109–115.
- Sara Tonelli and Rodolfo Delmonte. 2010. Venses++: Adapting a deep semantic processing system to the identification of null instantiations. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 296–299, Uppsala, Sweden, July. Association for Computational Linguistics.
- Stephen Tratz and Eduard Hovy. 2010. A Taxonomy, Dataset, and Classifier for Automatic Noun Compound Interpretation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 678–687, Uppsala, Sweden, July. Association for Computational Linguistics.
- Naushad UzZaman, Hector Llorens, Leon Derczynski, James Allen, Marc Verhagen, and James Pustejovsky. 2013. Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. In *Second Joint Conference on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, pages 1–9, Atlanta, Georgia, USA, June. Association for Computational Linguistics.
- Greg Whittemore, Melissa Macpherson, and Greg Carlson. 1991. Event-building through role-filling and anaphora resolution. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 17–24, Berkeley, California, USA, June. Association for Computational Linguistics.
- Be N. Zapirain, Eneko Agirre, Lluís Màrquez, and Mihai Surdeanu. 2010. Improving Semantic Role Classification with Selectional Preferences. In *The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 373–376, Los Angeles, California, June. Association for Computational Linguistics.