

Utilizzare FrameNet

In quest'ultima sezione parliamo di come sia possibile utilizzare FrameNet per risolvere alcuni task classici.

FrameNet è una risorsa lessicale che mira a fornire, per una discreta porzione del vocabolario inglese, **un corpus di frasi annotate** sintatticamente e semanticamente.

Da questo corpus annotato è **possibile estrarre informazioni** sui modi in cui i termini vengono usati.

Finalità di FrameNet:

La finalità del progetto è di fornire un inventario di situazioni ed eventi per abilitare la comprensione di un testo mediante tecniche di elaborazione del linguaggio naturale che siano in grado di associare automaticamente uno o più frame a ogni frase del testo stesso (disambiguazione di frame) insieme ai corrispondenti ruoli semantici (Gildea e Jurafsky, 2002).

Il task di information extraction consiste fondamentalmente nell'estrazione delle informazioni più rilevanti da parte di un documento al fine di crearne una versione ridotta.

Se per esempio ci indirizziamo sulla cronaca nera, operiamo una selezione sui termini che più frequentemente occorrono in tale contesto.

Anziché muoverci totalmente a caso sul documento senza avere alcuna idea del significato complessivo delle frasi che trattiamo, possiamo appoggiarci ad uno strumento come FrameNet al fine di operare più intelligentemente il nostro task.

Information extraction

I dati FrameNet devono dunque essere utilizzati per:

La WSD

La composizione semantica, cioè l'integrazione di informazioni associate con i dipendenti semantici in frames evocati dai loro reggenti semantici.

La scelta fra possibili -alternative- analisi sintattiche.

L'attivazione di un vocabolario topic-related per il riconoscimento e la selezione del senso nelle varie parti in cui si articola un testo.

Nota:

Supponiamo inoltre che esistano altre risorse per l'analisi sintattica, la risoluzione dell'anafora, ecc.

Articolo di cronaca nera

Consideriamo dunque le seguenti frasi che costituiscono un articolo di cronaca nera:

1. Washington (CNN)- Alleged White House gunman Robert Pickett was **arraigned** Wednesday at a federal court in Washington and ordered held without bond.
2. A federal magistrate informed Pickett of the charges against him - assaulting a federal officer with a deadly weapon, which carries a maximum of ten years in prison.
3. The magistrate set a preliminary hearing for next Tuesday and ordered Pickett held without bond.
4. Pickett, who was shot in the knee by the Secret Service after allegedly firing two shots outside the White House, used crutches to walk into the court.
5. He did not enter a plea.

Come vediamo, la prima frase informa dell'arresto. Nella seconda l'accusato è informato delle accuse che gli vengono mosse. Poi viene data l'informazione sulla data per l'udienza preliminare, ecc.

Come
procedere

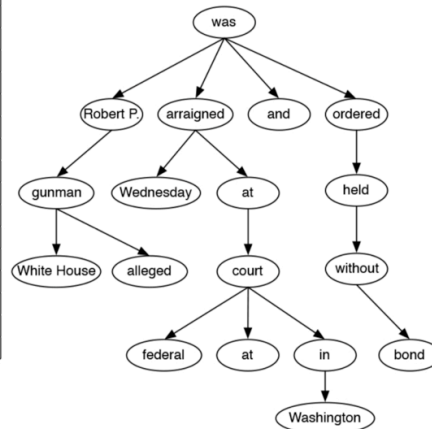
Il nostro intento è ora quello di, in seguito ad una **analisi sintattica**, cercare di associare un certo numero di frame alle frasi e di annotarne le componenti principali.

A priori, non possiamo sapere di quanti frame consti la situazione descritta in un articolo pertanto proviamo ad associare un frame diverso ad ogni frase.

Un primo modo per vedere come gli elementi semanticamente dipendenti sono inseriti nella struttura del frame, è analizzare le relazioni (di natura *sintattica*) fra reggenti e dipendenti

Index	Dependents	Form	POS
8	none	Alleged	adj
7	none	White House	n
6	7,8	gunman	n
2	[6]	Robert Pickett	n
1	2,3,4,5	was	be
3	(2)9,10	arraigned	ppt
9	none	Wednesday	adv
10	11	at	prp
13	none	a	art
12	none	federal	adj
11	12,13,14	court	n
14	15	in	prp
15	none	Washington	n
4	none	and	cnj
5	(2)16	ordered	ppt
16	(2)17	held	ppt
17	18	without	prp
18	none	bond.	n

1. Washington (CNN)- Alleged White House gunman Robert Pickett was arraigned Wednesday at a federal court in Washington and ordered held without bond.



Il frame di Arraignment

In Figura vediamo l'albero sintattico per la prima frase del nostro articolo.

Il modo più semplice di utilizzare FrameNet consiste nello scegliere una parola (a partire da quella semanticamente più rilevante) per capire quali frame possono essere attivati da quella parola.

Una volta identificati i ruoli semantici (i FE), si cerca di accoppiare le varie "necessità semantiche" di ogni frame attivato (i.e., i vari elementi -FE- del frame attivato), con le porzioni della frase in input. Se vi sono più frame attivati dalla stessa frase, il frame che meglio si adatta alla proposizione è quello scelto.

Il frame Arraignment

Nel caso specifico della prima frase, notiamo che i dettagli temporali del momento dell'arresto devono essere calcolati in base alla data dell'articolo (*Wednesday, next Tuesday*) e che devono essere risolti i riferimenti anaforici fra il *R. Pickett* della prima frase e i vari successivi riferimenti.

Per la prima frase il frame attivato è quello dell'**Arraignment** (citazione in giudizio).

In tale contesto, una delle funzioni di una corte è quella di informare l'imputato di quali accuse gli sono mosse. Pertanto, *federal magistrate, inform, charges against* possono essere incorporate nella rappresentazione semantica del testo.

La bellezza di FrameNet consiste nel fatto che, se si trova il frame corretto, i problemi di ambiguità sono tendenzialmente risolti. In particolare:

Disambiguazione

I significati di *magistrate* come ufficiale giudiziario (invece che capo della magistratura) e di *charge* (come accusa, invece che come carica) possono essere determinati per mezzo di un criterio di coerenza (dettato dal frame scelto) e non sono pertanto ambigui.

Blocchi di parole

Osserviamo che l'analisi non può essere limitata ad un approccio parola per parola, ma entità come *White House, Robert Pickett, e the Secret Service* devono essere considerate in toto.

Inoltre, lo stretto legame fra le parole deve essere riconosciuto per i seguenti sintagmi: *held without bond, assaulting a federal officer with a deadly weapon, preliminary hearing, firing shots, enter a plea.*

Il frame Evidentiality

Le parole *alleged* (dichiarare/affermare) e *allegedly* (presumibilmente) sono casi particolari e vengono trattati come segue:

Quando l'aggettivo presunto precede un costituente nominale è associato con la categoria di tale nome;

Quando *allegedly* (i.e., "secondo quanto si dice") si riferisce al verbo reggente, significa che si tratta di una notizia riportata da altra fonte, e funziona da liberatoria per il giornalista che in tal modo attribuisce ad altri la paternità di tale affermazione;

Però, quando *allegedly* è riferito a un verbo che descrive un subevento, (*who was shot in the knee by the Secret Service after allegedly firing two shot outside the White House*) nasce un problema sullo scope dell'avverbio. Si tratta in questo caso di uno strumento con cui l'autore di un articolo aggiunge qualche tipo di informazione testimoniale a una parte della descrizione.

Il frame Evidentiality, che viene evocato in queste situazioni, è un *frame metalinguistico* che ha come FE l'Ascriber (colui che attribuisce) e la Description.

Il vocabolario di tale frame comprende *alleged(ly)*, *reported(ly)*, *known*, *certified*, *certifiably*, *authentic*, *suspected*, *self-described*, *admitted* e *pochi altri*.

I vari frame evocati dall'articolo di cronaca nera che stiamo trattando:

Frame evocati
nell'esempio

Frame	Elaborates	Elements	Sub-Frames
Court		Judge(s) Officer(s) Courtroom	
Criminal Process		Court Prosecution (group) Prosecutor Defense (group) Defendant(s) Defense Attorney(s) Charges Law	Court Appearance Arrest Accusation Arraignment Preliminary hearing Trial Verdict
Arraignment		Defendant Court	Stating of Charges Entering of Plea
Confinement		Jailer Prisoner	
Pre-trial Confinement	Confinement	Bail	Confinement Posting of bail Release on bail Return to court Flight
Appointment		Prot1 Prot2	Promise Action
Court-date-setting	Appointment	Judge = Prot1 Defense & Prosecution = Prot2 Action = subframe of criminal process Action.Place = courtroom	
Event		Theme ("affected object") (Cause) (Result) Place Time	
Action	Event	Actor (Means) (Manner)	Action = Event
Action-Intentional	Action	Actor.type = Sentient	Forming of Intention Action = Action Result = Result
Action on Bodily target	Action	Prot1 = Actor Prot2 Prot2_body_part = Theme	

Alcuni dei frame evocati delle frasi dell'articolo e le relazioni fra loro

Conclusioni

In definitiva, FrameNet è una risorsa anglosassone per anglosassoni costruita manualmente e che permette di associare ad una data situazione (ovviamente descritta da un testo) un insieme di frame grazie ai quali dedurre informazioni semantiche sui vari componenti della frase associati ai relativi FE.

All'url <https://framenet2.icsi.berkeley.edu/> è possibile consultare la lista dei frame di FrameNet.