

FrameNet

FrameNet è un progetto svolto presso l'International Computer Science Institute a Berkeley, che consiste nella produzione di una risorsa elettronica basata sui cosiddetti **frame semantici**.

Un frame semantico fornisce una codifica schematizzata di una situazione, uno stato o un evento con l'ausilio di **unità lessicali** (ad esempio, per il frame compravendita, le parole comprare, vendere, costo, ecc.) e di **ruoli semantici** specifici per il frame (ad esempio, compratore, venditore, denaro, bene, ecc.).

L'assunzione dietro a FrameNet è il fatto che gli umani non percepiscano la realtà "*ogni volta da capo*", ma bensì che in qualche modo ogni persona abbia già un riferimento relativo a una data situazione/oggetto e che usi quello stereotipo per processare la realtà. Questa intuizione (che dobbiamo a Minsky) per quanto possa sembrare banale ha avuto enormi conseguenze sulle scienze cognitive.

Ciò che vogliamo fare con FrameNet è dunque mettere in relazione le parole col loro significato (un task del tutto nuovo eh?) usando però i frame che le rappresentano.

Il nucleo di FrameNet

Se ne deduce che il cuore di FrameNet siano proprio i frame e le attività che li riguardano; più precisamente:

Caratterizzare i frame (quali sono gli elementi significativi che caratterizzano il frame?)

Una volta definiti i ruoli e le funzioni principali all'interno di un frame, trovare le parole che meglio li rappresentino.

Per ogni frame vengono costruire delle frasi di esempio che devono essere rappresentative dei termini che abbiamo individuato ai passi precedenti.

Dopo aver ottenuto un numero sufficiente di esempi li andiamo ad annotare (capiremo in seguito cosa s'intende).

Frames

Abbiamo già parlato dell'idea di frame diverse volte in questo volume; come sappiamo un frame è una definizione più o meno dettagliata di una situazione tipica, una stereotipazione di cose/stati/eventi.

I concetti che utilizziamo per descrivere un frame sono dunque relativi al determinato frame che stiamo trattando.

Sappiamo anche che i frame sono fondamentalmente una struttura dati costituita da slot che, a seconda del loro contenuto, definiscono il frame stesso (ricordiamo che il contenuto di uno slot può essere a sua volta un frame).

I frame sono organizzabili in strutture gerarchiche.

Secondo i creatori di FrameNet vi sono alcune parole (*i.e. verbi*) che sono dette *evocative di frame*, ovvero in grado di rimandare immediatamente ad un certo frame.

Es. ristorante ci potrebbe ricordare il frame di un cameriere che porta un ordine ad un tavolo.

L'evocazione di un frame può essere in qualche modo pensata come l'attivazione di una regola all'interno di un sistema esperto: ad una certa parola (fatto) corrisponde l'attivazione di un certo frame (regola). Chiaramente, nel caso di termini ambigui una parola potrebbe evocare diversi frame: vedremo come trattare il problema in seguito.

Lexical Units

Il nemico contro il quale siamo chiamati a combattere è sempre il solito: l'**ambiguità**.

I termini polisemici risultano problematici poiché, come abbiamo detto poco sopra, possono attivare frame diversi.

Pertanto, anziché lavorare con parole lavoriamo con **Lexical Units** (d'ora in poi LU) ossia coppie di parole col loro significato. Tanto per essere chiari, in WordNet LU differenti appartengono a synset diversi, così come in FrameNet LU differenti (tipicamente) appartengono a frame diversi.

A proposito di polisemia, possiamo notare che WordNet (così come FrameNet) sono risorse famose per la loro completezza nel catturare molte sfumature di significato (si pensi che il synset in cui si trova run contiene 16 significati) e questo può, dal punto di vista computazionale, rappresentare un problema (ad esempio nella risoluzione del task di WSD).

FrameNet è quindi molto ricca di LU e, talvolta, se ne effettua un clustering, lasciando esposti solo quelli più rilevanti (i.e., i più usati).

Problemi verbi

Dedichiamo ora qualche riga per discutere di alcune problematiche relative ai verbi; problematiche che nascono quando questi vengono utilizzati in contesti figurati ed in senso non letterale.

Si consideri che FrameNet riesce a gestire gran parte delle problematiche che stiamo per trattare (a differenza degli altri strumenti studiati fin'ora). Per fare ciò, viene compiuta un'analisi sintattica e a partire dal verbo si applicano dei pattern per cercare di comprendere in quale situazione ci troviamo.

In sostanza sfruttiamo indizi sintattici per comprendere il contesto corrente e quindi il contesto d'uso del verbo.

Stesso contesto
significato diverso

Consideriamo le frasi:

She earns a lot less than she deserves.

I made a lot of money, but I earned it.

Come vediamo, il verbo **earn** ha due significati diversi nelle due frasi appena riportate, seppure i due contesti siano del tutto simili.

Nella prima proposizione ci riferiamo ad un vero e proprio guadagno, mentre nella seconda frase ci riferiamo a un guadagno inteso come merito (*me li sono meritati*).

FrameNet ci può aiutare a risolvere questo tipo di ambiguità.

50 shade of
significato

Il verbo **replace** può essere inteso
in almeno tre modi diversi:

put (something) back where it belongs
(rimettere a posto)

occupy a position formerly occupied by (something/somebody),

put something a position formerly occupied by (something/somebody).

Gli ultimi due significati sono simili ma nell'ultima frase usiamo il verbo in senso transitivo.

FrameNet può aiutarci a discernere il senso mediante il quale un verbo è utilizzato.

Metonimie Subject as Medium

Tendenzialmente il soggetto di una frase che compie una determinata azione è la persona/oggetto che propriamente compie l'azione; si parla in questi casi di **Subject as Speaker** (e.g., Mom explained, you complained, ecc.).

Vi sono però casi in cui il soggetto non compie effettivamente l'azione indicata dal verbo, ma è soltanto un soggetto metaforico. In questi casi, si parla di **Subject as Medium** (e.g., Chapter 2 explained, your letter complained, ecc.).

Gli esempi *Subject as Medium* fanno dunque uso di metonimie e sono pertanto traducibili in frasi di più semplice interpretazione.

Es. *your letter complained... può diventare you complain in your letter that...*

FrameNet ci può aiutare in questo processo di traduzione.

Nota:

La metonimia è una figura retorica che consiste nella sostituzione di un termine con un altro che ha con il primo una relazione di vicinanza. La metonimia è una figura retorica del tutto paragonabile alla sineddoche ma mentre nella metonimia la relazione è di tipo qualitativo, nella sineddoche è di tipo quantitativo (la parte per il tutto, il genere per la specie, il singolare per il plurale).

Vi sono poi ulteriori usi più o meno particolari dei verbi che possono avere un uso cognitivo abbastanza importante e che quindi possiamo cercare di cogliere grazie a FrameNet.

Es. la proposizione *the heavy winds explain the number of winmills around here* afferma che il motivo per cui ci sono molti mulini è la grande ventosità ma ovviamente non possiamo sostenere che il vento spieghi davvero qualcosa.

Omissibilità

Consideriamo le frasi:

Do you want to meet the Red Cross representative? I already *gave*.

Did you remember a present for your daughter's birthday? I already *gave*.

Nel primo caso il verbo *give* è inteso come dare un contributo, mentre nel secondo caso si intende dare come consegnare.

Giacché con FrameNet noi vogliamo comprendere la situazione, queste sfumature di significato sono importanti.

Normalizzazione

Un ulteriore problema è la nominalizzazione, ossia, quando il verbo cambia significato a seconda del sostantivo con cui appare

Anche in questo caso, FrameNet può aiutarci a comprendere appieno il significato di un verbo.

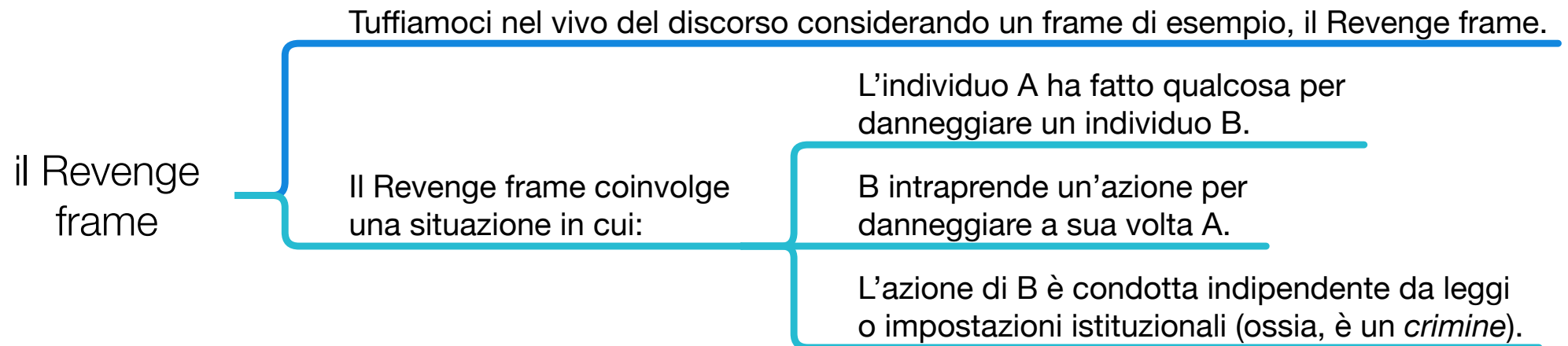
Es.

aderire alla pelle; (*aderenza*)

aderire ad un partito (*fedeltà/adesione*)

osservare una regola (*osservanza/rispetto*)

osservare un processo (*orsservazione*)



Vocabolario *rev frame*

All'interno del Revenge frame definiamo il seguente vocabolario (in realtà sono elementi LU: termine + significato):

Nomi: revenge, vengeance, reprisal, retaliation, retribution.

Verbi: avenge, revenge, retaliate (against), get back (at), get even (with), pay back.

Aggettivi: vengeful, vindictive.

Locuzioni verbo+nome: take revenge, exact retribution, wreak vengeance.

Notiamo due cose interessanti:

Abbiamo aggiunto ad alcuni verbi delle preposizioni al fine di discriminare i vari significati dei verbi stessi

Abbiamo incluso delle locuzioni solitamente adottate quando ci si trova nella situazione che il frame vuole descrivere.

Frame Elements *rev frame*

Definiamo poi i frame elements, ovvero le componenti vere e proprie del frame (in pratica gli elementi che hanno ruolo semantico) che andremo ad usare come annotatori in fase d'uso del sistema.

Forniamo inoltre una descrizione testuale del frame.

Ecco quindi che il nostro Revenge frame è formato da due parti:

1. Frame definition:

because of some injury to something-or-someone important to an avenger (maybe himself), the avenger inflicts a punishment on the offender. The offender is the person responsible for the injury.

2. Lista dei FE:

avenger

offender

injury

injured_party

punishment

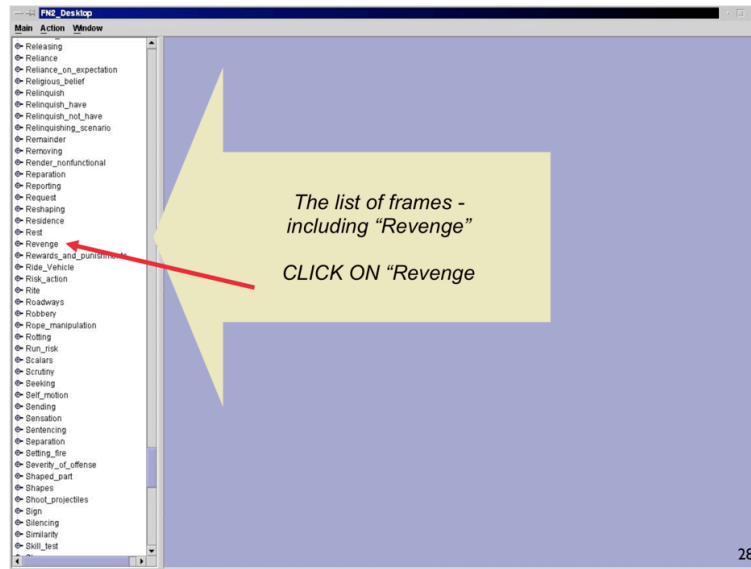
Nota:

Si noti che i FE scelti sono molto precisi e il meno generico possibili (non usiamo semantic role come *agent*, *patient*, *goal*, *etc.*), proprio per definire un frame che sia il più possibile aderente alla situazione che vogliamo descrivere.

Collezionare
gli esempi

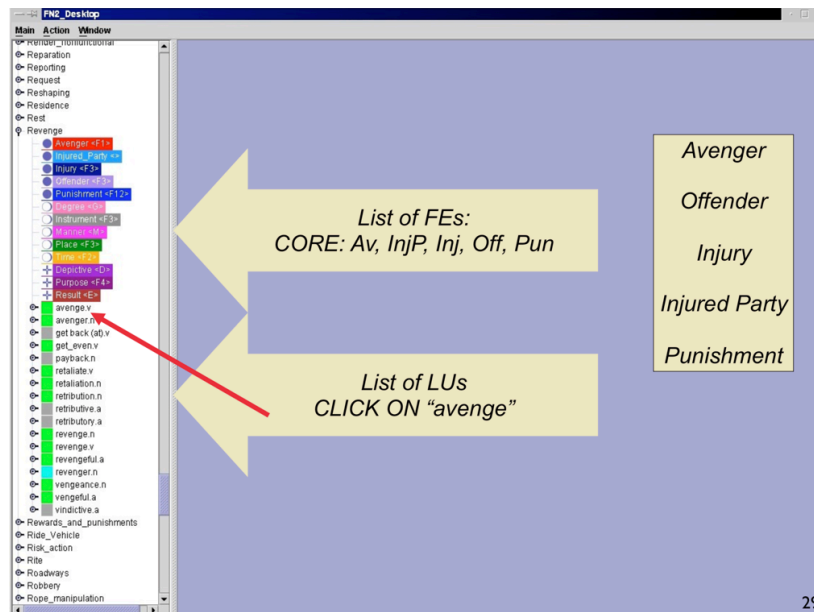
In questa sezione comprenderemo come FrameNet è stato popolato o osserveremo alcuni screenshot del sistema di backend per la gestione (e quindi la popolazione) di FrameNet.

Per popolare i frame di FrameNet è stato importato nel sistema un grande corpus (main corpus: *British National Corpus*) di esempi annotati (circa 200 milioni di parole) al fine di permettere agli esperti linguisti di rimappare ogni frase nel suo frame corretto, mostrando l'uso di ogni parola (FE) che lo compone



All'apertura del software abbiamo sulla sinistra una lista di frame: selezioniamo il nostro frame di interesse, il Revenge frame.

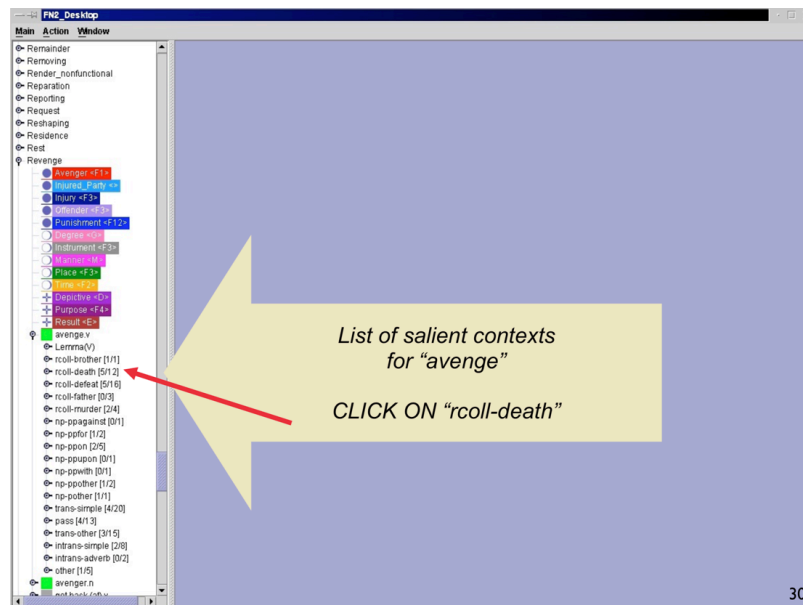
Software
FrameNet



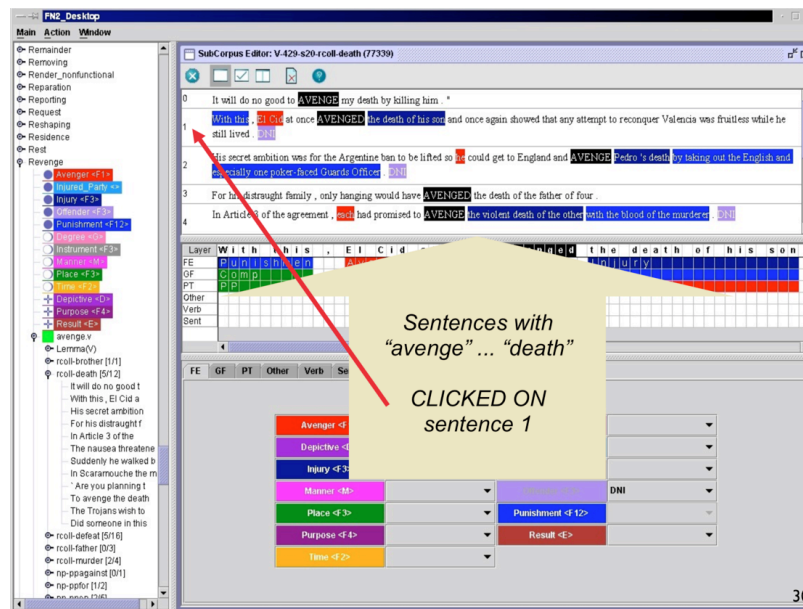
Possiamo osservare la presenza dei vari FE e anche dei LU.

Alcuni FE sono indicati con un pallino pieno, indicando il fatto che sono necessari per identificare il frame, i.e., se in una frase non si riesce ad identificare almeno uno dei FE obbligatori, la frase non viene considerata facente parte del frame.

Gli LU sono indicati mediante una notazione term.pos (e.g., avenger.v).

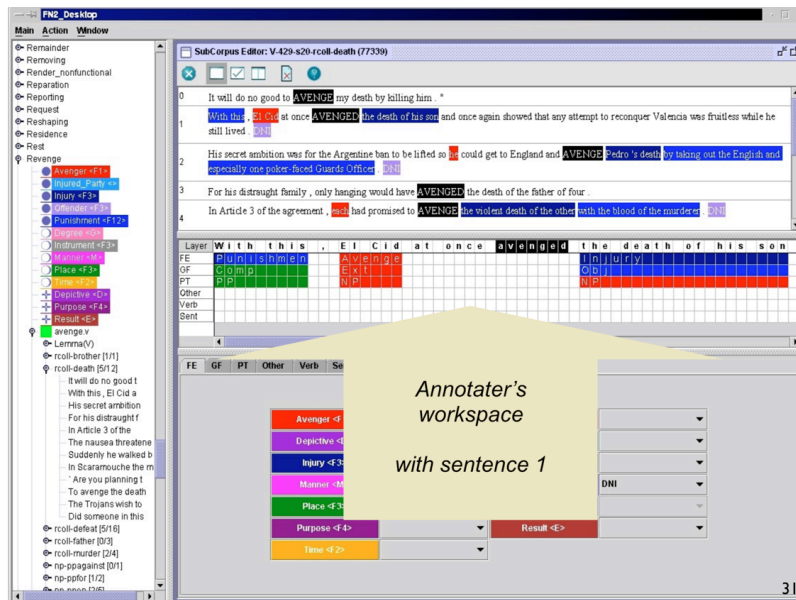


In questa figura vediamo il fatto che per ogni LU è possibile osservare i contesti di interesse in cui quel termine appare. *avenge.v* ad esempio è in relazione con defeat.



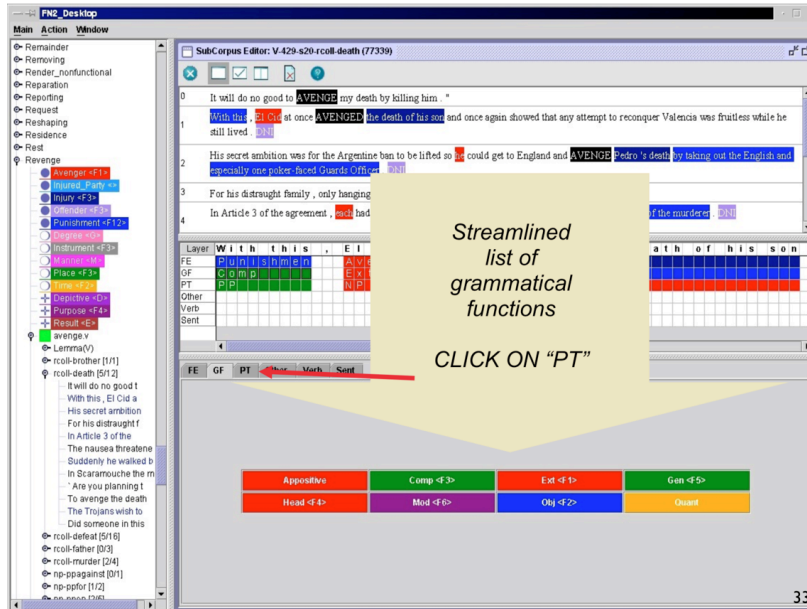
Software
FrameNet

Nella parte centrale del software possiamo vedere le varie frasi che sono state etichettate sotto il Revenge frame. L'etichettatura è esplicitata dalla concordanza di colori fra i token della frase e la lista degli FE a sinistra.



Lo spazio dell'annotatore, che può per l'appunto etichettare ogni parola/locuzione con un FE, una GF (Grammatical Function) e un PoS tag.

Il verbo non è annotabile poiché è il punto di partenza da cui stiamo costruendo il frame (il verbo evoca il frame).



Le funzioni grammaticali disponibili all'annotatore per il Revenge frame

Software
FrameNet

Frequency data
e
project outreach
(*estensione*)

Facciamo presente il fatto che i dati relativi alle frequenze di frame e di frasi non sono incluse in FrameNet al momento. D'altronde, questi dati varieranno di molto a seconda del tipo di corpus che si sta trattando.

In conclusione, aggiungiamo che sono stati fatti diversi tentativi per ottenere FrameNet multilingua, ma la scarsità di risorse ha spesso portato alla morte di questi progetti.