






Un Sistema per la Redazione Assistita di Testi Giuridici

Tesi di Laurea Specialistica in Informatica di
Emiliano Giovannetti





Strumenti informatici per la gestione di dati testuali

tra gli obiettivi che ci si pone:

- recuperare la conoscenza codificata nei testi mediante tecniche in grado di estrarre ed elaborare il contenuto:

"contenuto passivo" → "contenuto attivo"

testo significato del testo





L'elaborazione del contenuto: analisi e generazione di testi

elaborazione del contenuto

Analisi Generazione

estrazione del contenuto da un testo creazione di testo e relativa rappresentazione del contenuto






Una metodologia per la generazione assistita di testi

Combina tecniche per:

- l'elaborazione del linguaggio naturale
- la rappresentazione della conoscenza

al duplice scopo di:

- generare un documento testuale
- generarne la rappresentazione del contenuto





Il dominio di applicazione: documenti legali

La **legislazione** è una tra le fonti del **diritto** che più necessita di strumenti di supporto automatici:

- la **legimatica** studia le tecniche per la redazione di documenti legali

Dominio di applicazione di questa ricerca: strumenti per generare **contratti di lavoro**.

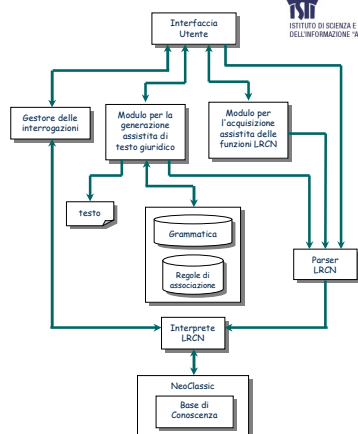


Obiettivi della legimatica

- redazione di testi chiari e non ambigui:
 - l'articolo 169 del Codice Civile Italiano è costituito da un'unica frase in cui compaiono tre periodi ipotetici negativi, di cui uno formulato in modo passivo
 - "Se non è stato espressamente consentito nell'atto di costituzione, non si possono alienare, ipotecare, dare in pegno o comunque vincolare beni del fondo patrimoniale se non con il consenso di entrambi i coniugi e, se vi sono figli minori, con l'autorizzazione concessa dal giudice, con provvedimento emesso in camera di consiglio, nei soli casi di necessità o di utilità evidente." (C.C. 169)
- redazione di testi giuridici anche da parte di utenti inesperti di tecnica legislativa.



Architettura del sistema



Un linguaggio per la rappresentazione di concetti normativi (LRCN)

- definizione di un **linguaggio** in grado di rappresentare "concetti normativi", ovvero:
 - modalità utilizzate dal legislatore per organizzare la conoscenza del testo

proposizione (testo) → LRCN → concetto (significato)








Concetti normativi ricorrenti all'interno di un contratto



- Definizione**
 - per generalizzazione**
 - "Il contratto di fornitura di lavoro temporaneo è un tipo di contratto."*
 - per aggregazione**
 - "La verifica delle condizioni di lavoro avviene all'inizio dell'attività."*
- Implicazione**
 - "se il dipendente timbra il cartellino il turno di lavoro comincia"*
- Attribuzione di un diritto**
 - "il lavoratore ha il diritto di scioperare"*
- Attribuzione di un dovere**
 - "l'ente ha il dovere di retribuire il dipendente"*




Funzioni LRCN

Concetti normativi	LRCN	Semantica
definizione per generalizzazione	defGen(entità, entità più generale)	codice NeoClassic
definizione per aggregazione	defAgg(entità, verbo, tipo complemento, oggetto del verbo, annesso)	codice NeoClassic
attribuzione di un diritto a qualcosa	dirittoEntità(entità, entità di diritto, annesso)	codice NeoClassic
attribuzione di un diritto a fare qualcosa	dirittoAzione(entità, verbo, tipo compl., oggetto del verbo, annesso)	codice NeoClassic
attribuzione di un dovere	dovere(entità, verbo, tipo compl., oggetto del verbo, annesso)	codice NeoClassic
implicazione	implicazione(fatto antecedente, fatto conseguente)	codice NeoClassic

La definizione per generalizzazione

- testo:**
 - "Il contratto di formazione è un tipo di contratto."*
- funzione LRCN**
 - defGen**("contratto di formazione", "contratto")
- interpretazione della funzione defGen**



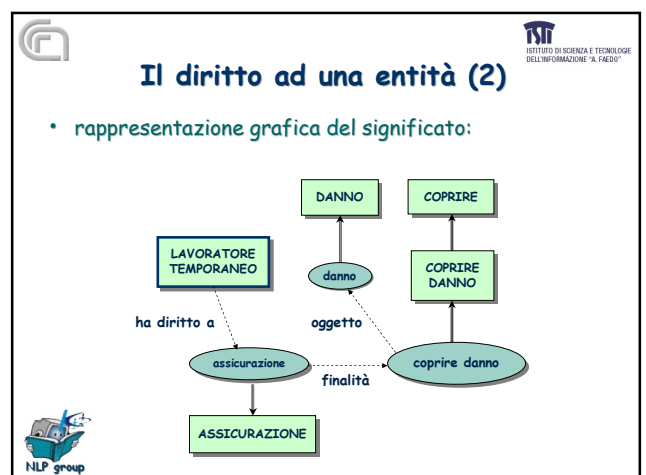
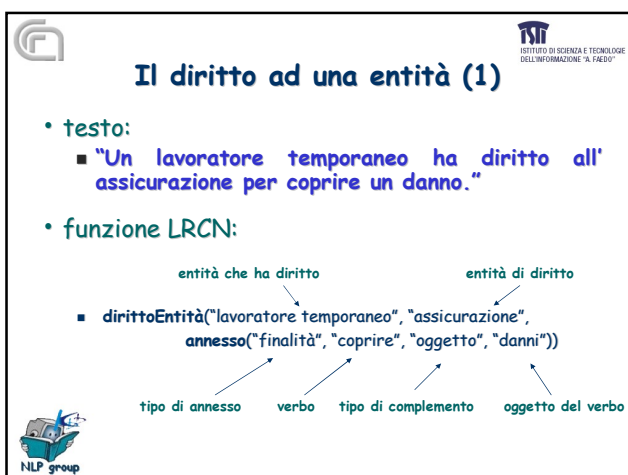
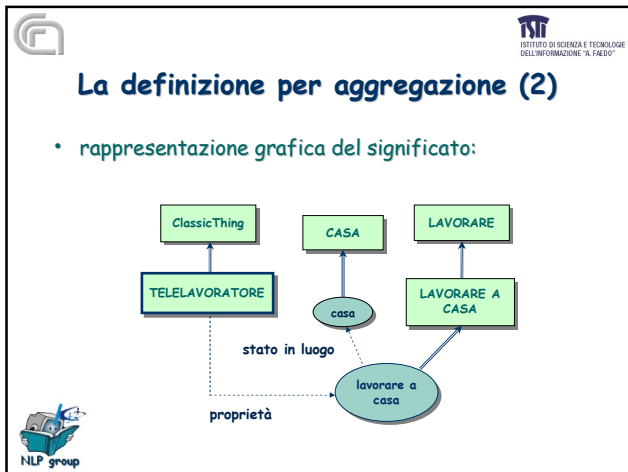



La definizione per aggregazione (1)

- testo:**
 - "Il telelavoratore lavora a casa."*
- funzione LRCN:**
 - defAgg**("telelavoratore", "lavora", "stato in luogo", "casa", nessunAnnesso)







Il diritto ad una azione (1)

- testo:
 - "Gli enti possono definire progetti di sperimentazione."
- funzione LRCN:
 - entità che ha diritto verbo tipo di complemento
 - `dirittoAzione("ente", "definire", "oggetto", "progetti di sperimentazione", nessunAnnesso)`
 - oggetto del verbo annesso

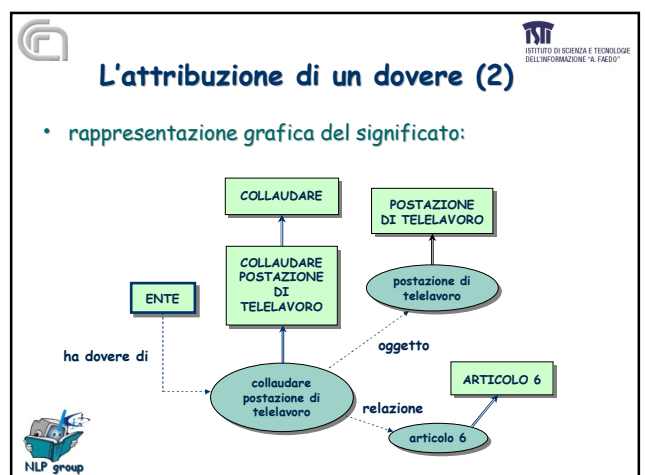
NLP group





L'attribuzione di un dovere (1)

- testo:
 - "L'ente ha il dovere di collaudare la postazione di telelavoro in conformità all'articolo 6."
- funzione LRCN:
 - entità che ha dovere verbo tipo di complemento oggetto del verbo
 - `dovere("ente", "collaudare", "oggetto", "postazione di telelavoro", annesso("relazione", "articolo 6", "", ""))`
 - tipo di annesso entità in relazione

NLP group




L'implicazione (1)



In una implicazione sono coinvolti due *fatti*:

- un **antecedente**
- un **conseguente**

Con un *fatto* è possibile modellare:

- avvenimenti,
- attribuzioni di proprietà, diritti e doveri *temporanei*



L'implicazione (2)

- **testo:**
 - "se il contratto scade il lavoratore ha diritto alla liquidazione."
- **funzione LRCN:**

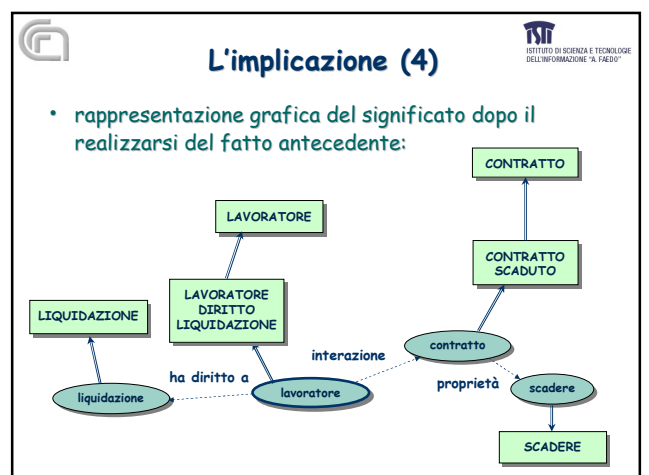
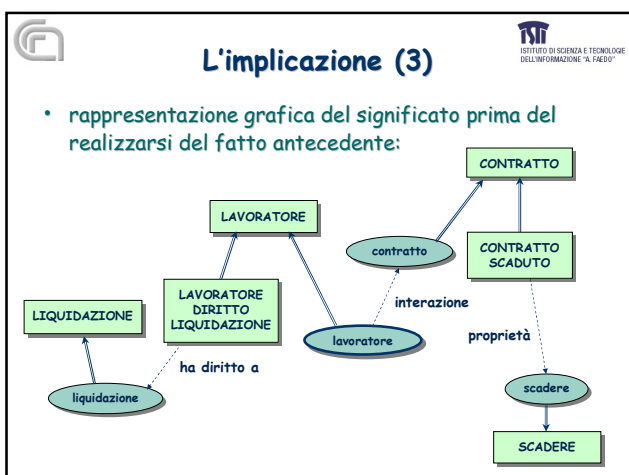
```

implicazione(
  fatto("contratto", "scade", " ", " ",
    nessunAnnesso),
  fattoDirittoEntità("lavoratore", "liquidazione",
    nessunAnnesso ));

```

Diagram labels:

- fatto antecedente (f.a.)
- soggetto del f.a.
- verbo del f.a.
- fatto conseguente (f.c.)
- soggetto del f.c.
- entità di diritto temporaneo





Il processo di generazione

• Generazione assistita del testo di un contratto

■ obiettivi:

- generazione del documento
- rappresentazione attiva del contenuto
 - dal testo alla "forma logica" in LRCN

■ caratteristiche:

- guida basata sulla grammatica del testo:
 - concetti normativi
 - forma linguistica e struttura testuale



La grammatica (1)

Produzioni relative alla struttura del contratto:

<Contratto> → <Titolo>
<Titolo> → <Capo>
<Capo> → <Articolo>
<Articolo> → <Comma>

Produzioni relative alla composizione di un comma:

<Comma> → <Definizione> <Altre strutture>* |
<Diritto> <Altre strutture>* |
<Dovere> <Altre strutture>* |
<Implicazione> <Altre strutture>*

<Altre strutture> → <Definizione> |
<Diritto> |
<Dovere> |
<Implicazione>



La grammatica (2)

Produzioni relative alla sintassi del testo:

<SN> → <N> | <Art> <N>
<Art> → il | la | l'
<N> → telelavoratore | lavoratore | utente | turno di lavoro
<SV> → <V> <SP>
<V> → lavora | timbra | comincia | assume
<SP> → <oggetto> | <stato in luogo> | <tempo>
<oggetto> → <SN>
<stato in luogo> → <prep. luogo> <SN>
<tempo> → <prep. tempo> <SN>
<prep. luogo> → a | in | nel
<prep. tempo> → il giorno | nel mese di | alle ore



La grammatica (3)

Produzioni relative ai concetti normativi:

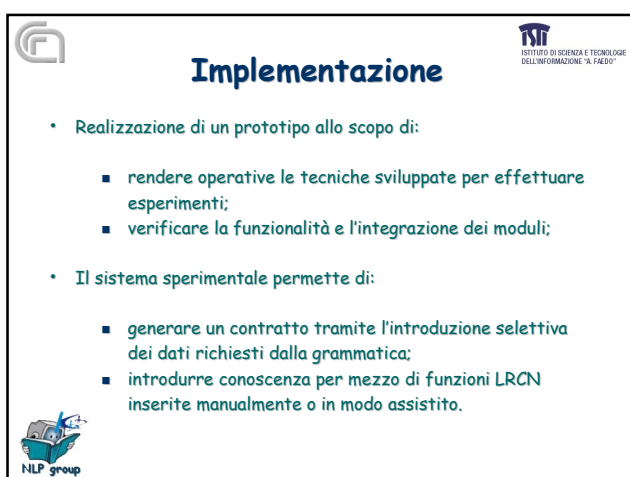
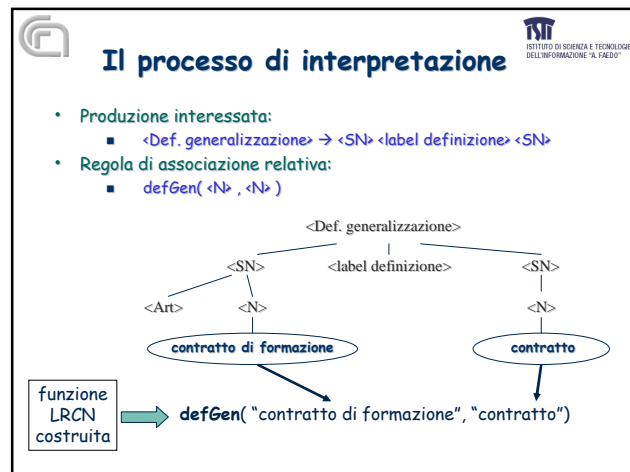
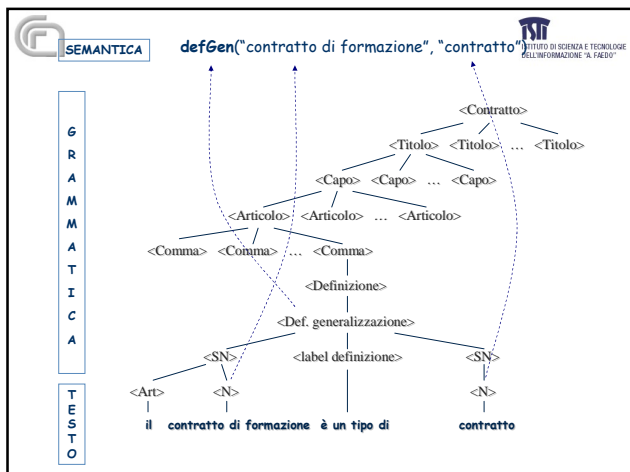
<Definizione> → <Def. aggregazione> | <Def. generalizzazione>
<Def. generalizzazione> → <SN> <label definizione> <SN>
<label definizione> → e' un | e' un tipo di

[...]

<Diritto> → <Diritto ad una entità> | <Diritto ad una azione>
<Diritto ad una entità> → <Diritto ad una entità senza annesso> |
<Diritto ad una entità annesso finalità> |
<Diritto ad una entità annesso relazione>

[...]

<Implicazione> → <Periodo ipotetico>
<Periodo ipotetico> → <P. condizionale> <P. principale>
<P. principale> → <Fatto>
<P. condizionale> → se <Fatto> | se <Fatto> allora





Una sessione d'esempio (2)

Produzione considerata relativa a <Contratto>:

<Contratto> → <Titolo>

Inserire il numero di <Titolo> da generare da 1 in su → 1

...

Le produzioni relative a <Comma> sono:

- 1) <Comma> → <Definizione> <Altre strutture>*
 - 2) <Comma> → <Diritto> <Altre strutture>*
 - 3) <Comma> → <Dovere> <Altre strutture>*
 - 4) <Comma> → <Implicazione> <Altre strutture>*
- Quale si vuole utilizzare? → 1



Una sessione d'esempio (3)

Le produzioni relative a <Definizione> sono:

- 1) <Definizione> → <Def. aggregazione>
 - 2) <Definizione> → <Def. generalizzazione>
- Quale si vuole utilizzare? → 2

Produzione considerata relativa a <Def. generalizzazione>:

<Def. generalizzazione> → <SN> <label definizione> <SN>

funzione LRCN aggiornata:

defGen(

Le produzioni relative a <SN> sono:

- 1) <SN> → <N>
 - 2) <SN> → <Art> <N>
- Quale si vuole utilizzare? → 2



Una sessione d'esempio (4)

Le produzioni relative a <Art> sono:

- 1) <Art> → il
- 2) <Art> → la
- 3) <Art> → l'

Quale si vuole utilizzare? → 1

TESTO GENERATO: *il*

Le produzioni relative a <N> sono:

- 1) <N> → ?
- 2) <N> → lavoratore
- 3) <N> → telelavoratore
- 4) <N> → cartellino
- 5) <N> → turno di lavoro

Quale si vuole utilizzare? → 3



Una sessione d'esempio (5)

TESTO GENERATO: *il telelavoratore*

funzione LRCN aggiornata:

defGen(telelavoratore,

Le produzioni relative a <label definizione> sono:

- 1) <label definizione> → è un
 - 2) <label definizione> → è un tipo di
- Quale si vuole utilizzare? → 2

TESTO GENERATO: *il telelavoratore è un tipo di*

Le produzioni relative a <SN> sono:

- 1) <SN> → <N>
 - 2) <SN> → <Art> <N>
- Quale si vuole utilizzare? → 1



Una sessione d'esempio (6)

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"



Le produzioni relative a <N> sono:

- 1) <N> → ?
- 2) <N> → lavoratore
- 3) <N> → telelavoratore
- 4) <N> → cartellino
- 5) <N> → turno di lavoro

Quale si vuole utilizzare? → 2

TESTO GENERATO: *il telelavoratore è un tipo di lavoratore*

funzione LRCN aggiornata:
defGen(telelavoratore, lavoratore);







Vantaggi del sistema presentato

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"



- redazione di testi chiari e non ambigui;
- redazione di testi giuridici anche da parte di utenti inesperti di tecnica legislativa;


funzionalità aggiuntive ↓

- verifica in fase di redazione della consistenza del contenuto del testo;
- accesso alla rappresentazione, durante e dopo la redazione, per interrogazioni e modifiche;
- esecuzione di processi inferenziali, anche in fase di redazione, in grado di fornire nuova conoscenza.



Sviluppi futuri

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"



- si può prevedere una integrazione con strumenti linguistici:
 - dizionari, thesauri, ecc.
- la base di conoscenza rappresentata in un formalismo di nuova generazione:
 - OWL
- la tecnica presentata potrebbe essere "capovolta":
 - da generazione ad analisi (acquisizione della conoscenza)

