Generazione automatica di word cloud dinamiche

Ingegneria Informatica e dell'Automazione

Università degli Studi di Perugia FACOLTÀ DI INGEGNERIA

Candidato: Enrico Spataro Relatore: Prof ssa Carla Binucci Correlatore: Prof. Walter Didimo

19 02 2016



Contenuti

- 1 Introduzione
- 2 Creazione word cloud Word cloud statica Word cloud dinamica
- 3 Risultati sperimentali
- 4 Conclusioni



Obiettivi

Mostrare l'evoluzione di un testo tramite:

- creazione word cloud semantiche ad intervalli regolari;
- animazioni tramite morphing.



Cos'è una word cloud?



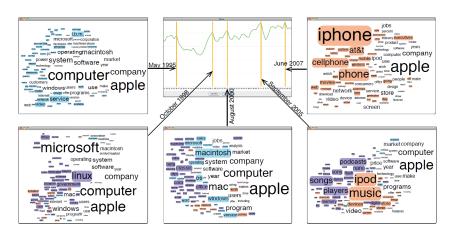


Word cloud semantiche

interrogator ESP kind mistakes identification guess Christmas Lovelace Babbage pupil fault clairvoyance telepathy criterion satisfy moment circumstances
Turing solipsist analogy infinite discrete-state **skin** really viva voce think inclined satisfactorily speed suitable induction child programme surprise fallacies suppose Yes conjectures storage random Machine Manchester propositions wheel imperatives wheel imperatives are recommended. pile neutron disabilities imitation think inclined game behaviour man argument quite clicks homework answer wrong try instance convenient packets store instruction obey computer



Word cloud dinamiche



Morphing





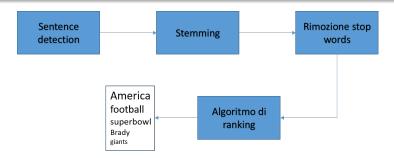
- Estrazione parole chiave
- Calcolo similarità
- 3 Disegno word cloud
- 4 Clustering



Generazione word cloud semantica

- Estrazione parole chiave

- 4 Clustering





Creazione word cloud

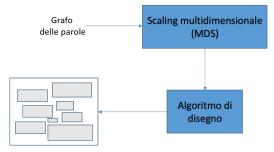
Word cloud statica

Generazione word cloud semantica

- Estrazione parole chiave
- 2 Calcolo similarità
- 4 Clustering



- Estrazione parole chiave
- 3 Disegno word cloud
- 4 Clustering





Generazione word cloud semantica

- Estrazione parole chiave
- 3 Disegno word cloud
- 4 Clustering





00000

Obiettivo: create K word cloud dal testo, si vuole passare da una word cloud alla successiva in modo graduale tramite tecniche di morphing.

ommario Introduzione **Creazione word cloud** Risultati sperimentali Conclusion

000000

Word cloud dinamica

Morphing

Caratteristiche

- Numero di frame
- Gestione stato delle parole
- Gestione del colore delle parole



Stato delle parole

Le parole, tra una word cloud e la successiva, possono:



Le parole, tra una word cloud e la successiva, possono:

- scomparire







Stato delle parole

Le parole, tra una word cloud e la successiva, possono:

- apparire



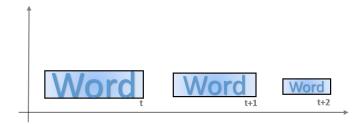




Stato delle parole

Le parole, tra una word cloud e la successiva, possono:

- rimanere nel layout (variando posizione)





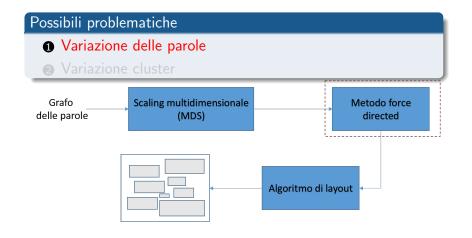
Generazione word cloud dinamica

000000

Possibili problematiche

- Variazione posizione delle parole
- 2 Variazione dei cluster





Generazione word cloud dinamica

000000

Possibili problematiche

- Variazione delle parole
- 2 Variazione cluster



Inserire video



Risultati sperimentali

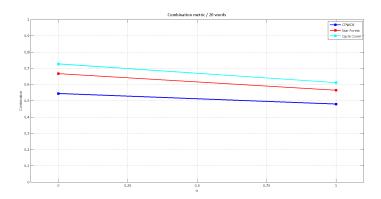
Setup:

- 200 discorsi, lunghezza media 17 minuti.
- 4 campionamenti per ogni testo.
- Parole estratte: 20,40,60

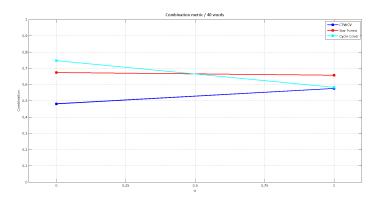
Metriche

- Combination metric
 - Distortion metric
 - Coherence metric
- Space metric
- Tempo d'esecuzione

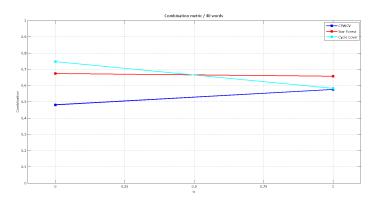




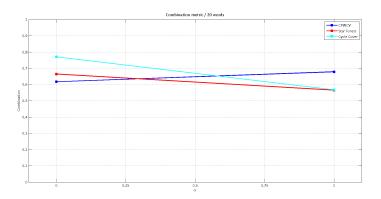




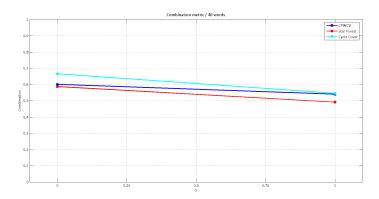




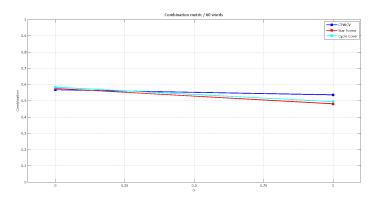




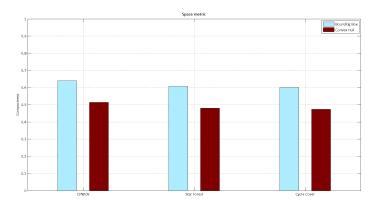








Space metric



Risultati sperimentali

Tempo d'esecuzione

- Configurazione caso peggiore: TFIDF, Jaccard Similarity e Star Forest
- Configurazione caso migliore: Term Frequency, Cosine Similarity e CPWCV

Parole estratte	Caso peggiore	Caso migliore
20	2.65	0.77
40	3.01	0.94
60	3.66	1.25

Concludendo...

Il lavoro di tesi si è articolato nelle seguenti fasi:

- Lorem ipsum
- 2 Lorem ipsum
- S Lorem ipsum

Lorem ipsum:

- Lorem ipsum
- Lorem ipsum

