

GOOGLE: COME LA MATEMATICA PUÒ FARE DIVENTARE RICCHI

Due studenti dell'università di Stanford, Sergey Brin e Larry Page, hanno fatto la loro fortuna inventando "Google".

Il problema di qualunque navigatore del web che sta effettuando una ricerca è quello di avere un elenco delle pagine web che gli possono interessare, **ordinate in base alla loro importanza**. Ma come si può definire l'importanza intrinseca di una pagina (detta tecnicamente *page rank*)?

Prima di Google, i motori di ricerca seguivano varie strade per ordinare le pagine trovate:

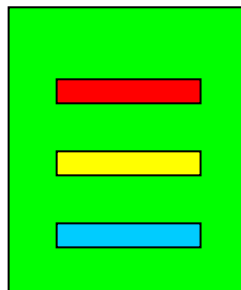
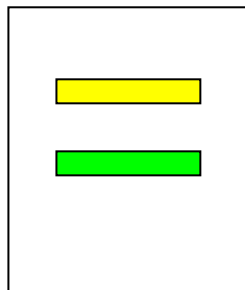
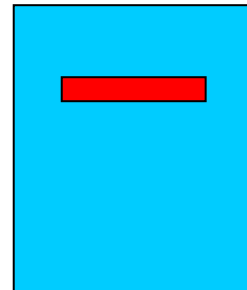
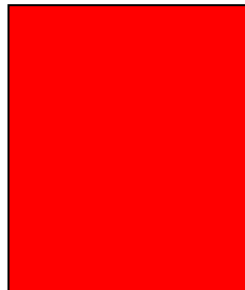
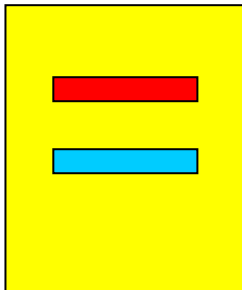
- in base al numero di volte che la parola cercata compare nella pagina
- in base al numero dei link che partono dalla pagina
- in base al numero dei link che arrivano alla pagina...

L'idea vincente di Brin e Page fu invece quella di definire l'importanza di una pagina **in base al numero delle pagine importanti che puntano ad essa**. Cercheremo, attraverso vari esempi, a dare un significato matematico a questa idea di base e da questo dedurre un metodo di calcolo.

ESEMPIO N.1

Le 5 pagine web disegnate in figura (ognuna contraddistinta da un colore) "puntano" con opportuni link alle altre. Ad esempio, la striscia azzurra nella pagina gialla indica che esiste un link dalla pagina gialla verso la pagina azzurra.

Qual è la pagina più importante? E la meno importante? Si può fare una classifica completa delle 5 pagine?



ESEMPIO N.2

La rete ferroviaria della Regione Lombardia è schematizzata in figura. Sono evidenziate le stazioni corrispondenti agli 11 capoluoghi di provincia, e i relativi collegamenti ferroviari diretti. Si può fare una classifica completa, in ordine di importanza, delle 11 stazioni? Trovate qualche analogia con l'esempio n.1 delle pagine web?

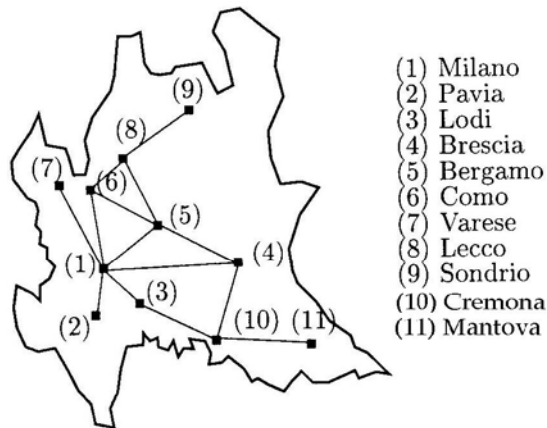


Fig. 6.2. Rappresentazione schematica delle connessioni ferroviarie per la sola Lombardia

ESEMPIO N.3

Abbiamo generato i dati relativi a 300 pagine web linkate a partire dalla home page dell'Università di Harvard (<http://www.harvard.edu>). La figura sottostante schematizza i collegamenti tra le pagine, che sono state opportunamente numerate: i rombi blu stanno ad indicare la presenza di un link (ad esempio, il rombo sulla 300-esima riga e sulla 22-esima colonna rappresenta un link dalla pagina n.22 alla pagina n.300). Il valore "nz" indica il numero totale dei link esistenti. Sarà possibile determinare l'importanza di ogni pagina in tempi ragionevoli?

