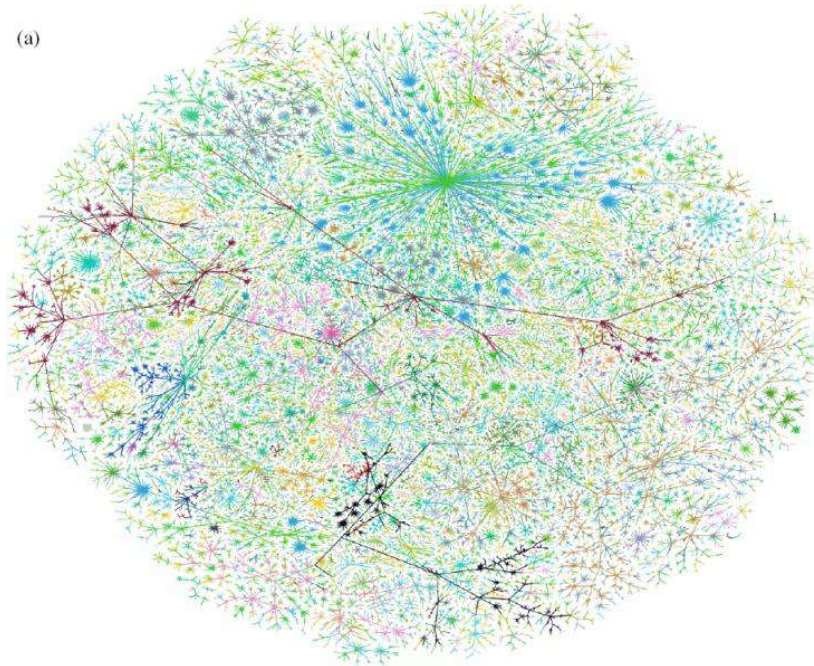


# RETI SOCIALI



## STRUTTURA DEL WEB

# Reti di Informazioni

- Fino ad ora ci siamo interessati soprattutto di reti sociali
  - I nodi della rete erano entità sociali che interagivano con altri soggetti e prendevano delle decisioni
  - I link della rete rappresentavano opportunità di relazioni sociali ed economiche tra coppie di entità
- Le **reti di informazioni** sono reti costituite da pezzi di informazione (documenti) collegati tra loro
  - Esiste un link tra due documenti se sono collegati
- Esempi di reti di informazione
  - Web (nostro principale argomento)

# Reti Sociali e Reti di Informazione

Le reti di informazioni differiscono dalle reti sociali perché

- I nodi non sono soggetti autonomi e non prendono decisioni
- I collegamenti sono unidirezionali
  - In un documento X si può far riferimento ad un documento Y
  - non è detto che Y faccia riferimento ad X

Le reti di informazioni mostrano numerose caratteristiche comuni con le reti sociali

- Componente gigante
- Un gran numero di percorsi brevi
- Misure di centralità

# Il World Wide Web

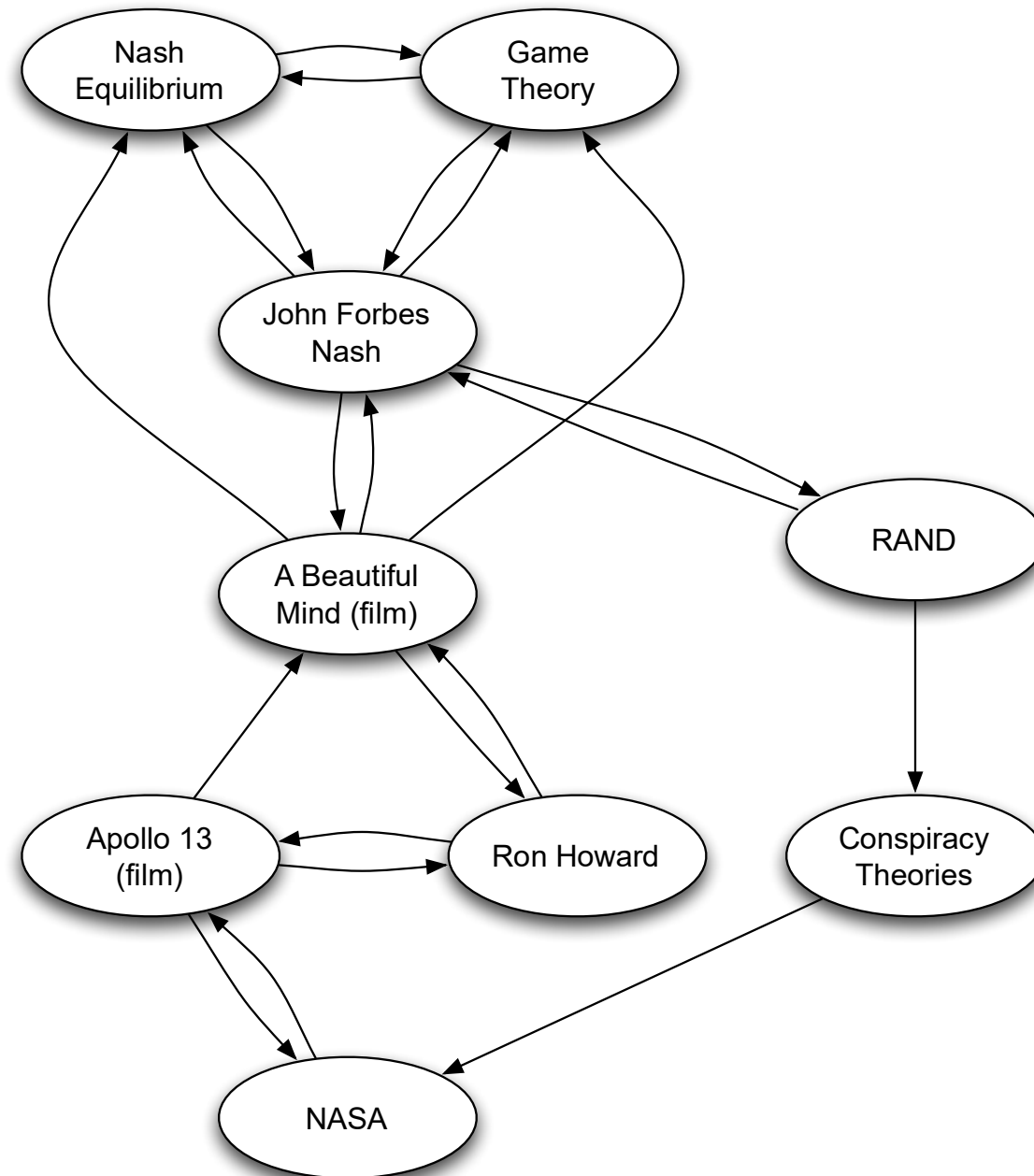
- Il Web è il più grande e popolare esempio di rete di informazioni
  - Creato da Tim Barners-Lee tra il 1989-1991 come strumento per rendere facilmente disponibili e fruibili i risultati dei vari gruppi di ricerca all'interno del CERN
- Idee base del progetto di Barners-Lee
  - Documenti realizzati in un formato testuale standard (pagine html) e resi disponibili in una zona del computer pubblicamente accessibile (web server)
  - un'applicazione (browser) specializzata per accedere alle pagine, visualizzarne il contenuto e navigarle
- Il Web oggi è un gigantesco sistema informativo costituito da miliardi di pagine
  - Le singole pagine sono realizzate da soggetti differenti che decidono autonomamente di renderle disponibili e le collegano ad altre pagine



# Iper testi

- Il Web è un'ipertesto
  - una pagina non deve necessariamente essere letta sequenzialmente
  - È possibile spostarsi da una pagina ad un'altra seguendo i collegamenti
    - I collegamenti sono predisposti dall'autore del testo
- Gli ipertesti sono stati introdotti a metà del XX secolo per rendere più fruibili i testi
- Precursori degli ipertesti
  - Reti di Citazioni e cross-reference tra libri ed articoli
    - Differiscono dai link dell'ipertesto perché un documento può citare solo i documenti creati prima di lui
  - Reti semantiche
    - i nodi rappresentano i concetti e gli archi rappresentano qualche tipo di relazione logica tra i concetti

# Reti di Riferimenti



# Vannevar Bush ed il Memex

- Gran parte delle caratteristiche attuali del Web sono state predette e descritte con straordinaria precisione nel 1945 da Vannevar Bush
  - As we May Think – Atlantic Monthly
- Bush immaginò che i futuri computer avrebbero sostituito i metodi tradizionali di memorizzazione lineare delle informazioni con meccanismi associativi
  - più simili a quelli del nostro cervello
- Bush immaginò un prototipo chiamato **Memex**
  - Molto simile al Web attuale
  - Conteneva versioni digitalizzate di tutta la conoscenza umana connesse da collegamenti associativi



# L'Evoluzione del Web

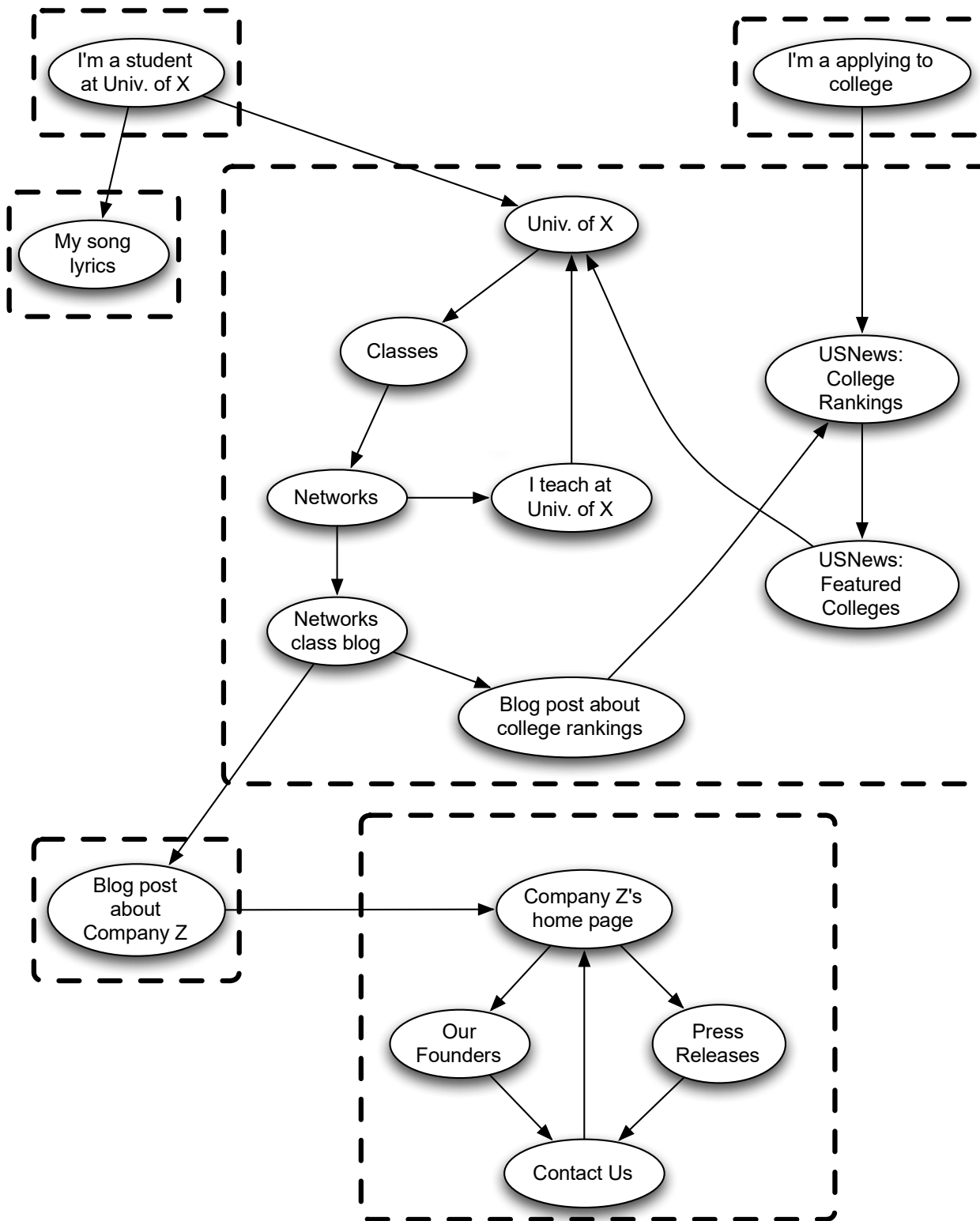
- Negli ultimi 25 anni il Web si è evoluto ed è diventato anche una grande piattaforma di calcolo
- Ad oggi le pagine web sono di due tipi
  - Pagine statiche che forniscono pezzi di informazione
  - Pagine dinamiche che presentano l'output di un programma eseguito sul server
- Allo stesso modo i link possono essere
  - Link navigazionali
    - Permettono di navigare tra le pagine statiche e di raggiungere le pagine dinamiche
    - Rappresentano la struttura del Web
  - Link transazionali
    - Permettono di lanciare l'esecuzione di un'applicazione
- I motori di ricerca prendono in considerazione solo le pagine raggiungibili tramite link navigazionali



# Il Web come Grafo Diretto

- Il Web e le reti di informazioni possono essere rappresentate come grafi diretti
  - Consideriamo solo i link navigazionali
- Il grafo del Web è
  - Diretto
  - Sconnesso
- Possiamo applicare al Web concetti di path, connettività e componenti connesse
- Una componente fortemente connessa  $S$  di un grafo diretto è un sottinsieme di nodi tali che
  - Ogni nodo di  $S$  è collegato (attraverso una qualche path) a tutti gli altri nodi dell'insieme
  - Il sottinsieme  $S$  è massimale

# Esempio



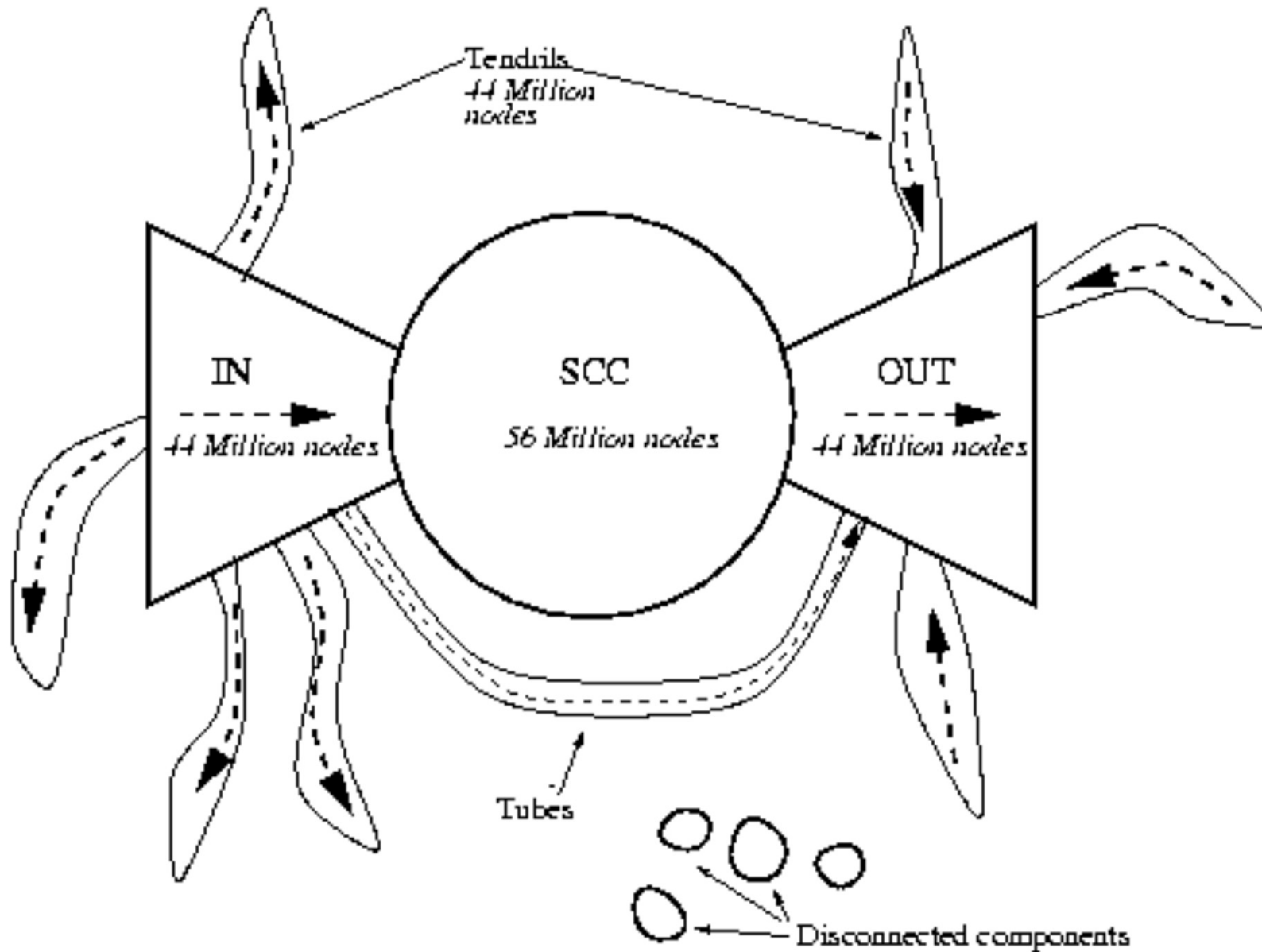
# La Struttura del Web

- Nel 1999 Broder ed altri hanno creato una mappa del Web del tempo
  - Basata sui dati immagazzinati nell'indice del motore di ricerca Altavista
- L'esperimento di Broder è stato replicato numerose volte, sia sul Web che su altre reti di informazioni
  - Google
  - Wikipedia
  - Reti di citazioni di articoli scientifici
  - Rete di prestiti interbancari
- Tutti gli esperimenti hanno mostrato che queste reti hanno strutture simili
  - Componente gigante

# Classificazione delle componenti del Web

- Il gruppo di Broder ha classificato le componenti fortemente connesse del Web in tre tipi
  - **CORE**: componente gigante in cui ogni pagina può raggiungere tutte le altre
  - **IN**: nodi che possono raggiungere il CORE ma non possono essere raggiunti da questo
  - **OUT**: nodi che possono essere raggiunti dal CORE ma non lo possono raggiungere
  - **TENDRILS**: nodi che possono essere raggiunti da IN o raggiungere OUT
  - **NODI ISOLATI**: tutto il resto
- Con la creazione di nuovi link pagine in IN e OUT passano in CORE e vengono sostituiti da nuove pagine appena create

# La Bow-Tie Structure



# L'emergenza del Web 2.0

- Negli ultimi 10 anni il contenuto informativo del Web si è arricchito enormemente
- Le principali trasformazioni sono state
  - Diffusione di siti che permettono a più persone di creare e mantenere contenuti condivisi (Wikipedia)
  - Sviluppo di applicazioni sociali che enfatizzano il collegamento tra le persone e non solo tra i documenti (facebook)
  - Trasferimento di moltissime informazioni personali online (google services, cloud)
- Tim 'O Reilly ha definito queste trasformazioni come l'emergenza del Web 2.0
  - Un'attitudine più che una nuova tecnologia
- I progettisti di siti web non possono prescindere dal considerare gli effetti di feedback sociale
  - Wisdom of crowds, long tail, cascading behaviours