



Il laboratorio

Costruzione del thesaurus, analisi a faccette, categorie

Manuela Grillo

Giornata di presentazione del Nuovo soggettario – Firenze, 8 Febbraio 2007

Costruzione del thesaurus, analisi a faccette, categorie

- Termini contenuti nel Prototipo
- Metodo di costruzione della gerarchia
- Profilo delle categorie e faccette
- Analisi a faccette
- Etichette di nodo e loro elaborazione
- Conseguenze della scelta della poligerarchia temperata sulle gerarchie preesistenti

Termini contenuti nel Prototipo

- **Termini della lista di partenza, o *Elenco di base***
 - a. tutti i termini degli *Aggiornamenti BNI 1956-1998* (pubblicati)
 - b. i nuovi termini introdotti dalla BNI dal 1999 al 2005 (mai pubblicati)
 - c. i termini di biblioteconomia, bibliografia, ecc., individuati nel *Soggettario e Aggiornamenti BNI*
 - d. termini provenienti dal *Soggettario* e dagli *Aggiornamenti BNI* ritenuti particolarmente obsoleti e bisognosi di revisione
- **Termini aggiunti in corso di elaborazione del Prototipo**
 - a. termini introdotti per la strutturazione dei termini dell'*Elenco di base*
 - b. termini derivanti dalla scomposizione di termini composti
 - c. termini impiegati nella sperimentazione della sintassi.



Metodo di costruzione della gerarchia

Due modalità di costruzione delle gerarchie, distinte e complementari, usate in combinazione:

- il **metodo induttivo** (dal basso verso l'alto)
- il **metodo deduttivo** (dall'alto verso il basso)

Prima fase: gerarchie costruite a partire da ciascun termine dell'Elenco di base, individuando i suoi termini sovraordinati fino al top term, con **metodo induttivo**. Parallelamente: preparazione di una **macrostruttura classificatoria**, costituita da categorie e faccette.

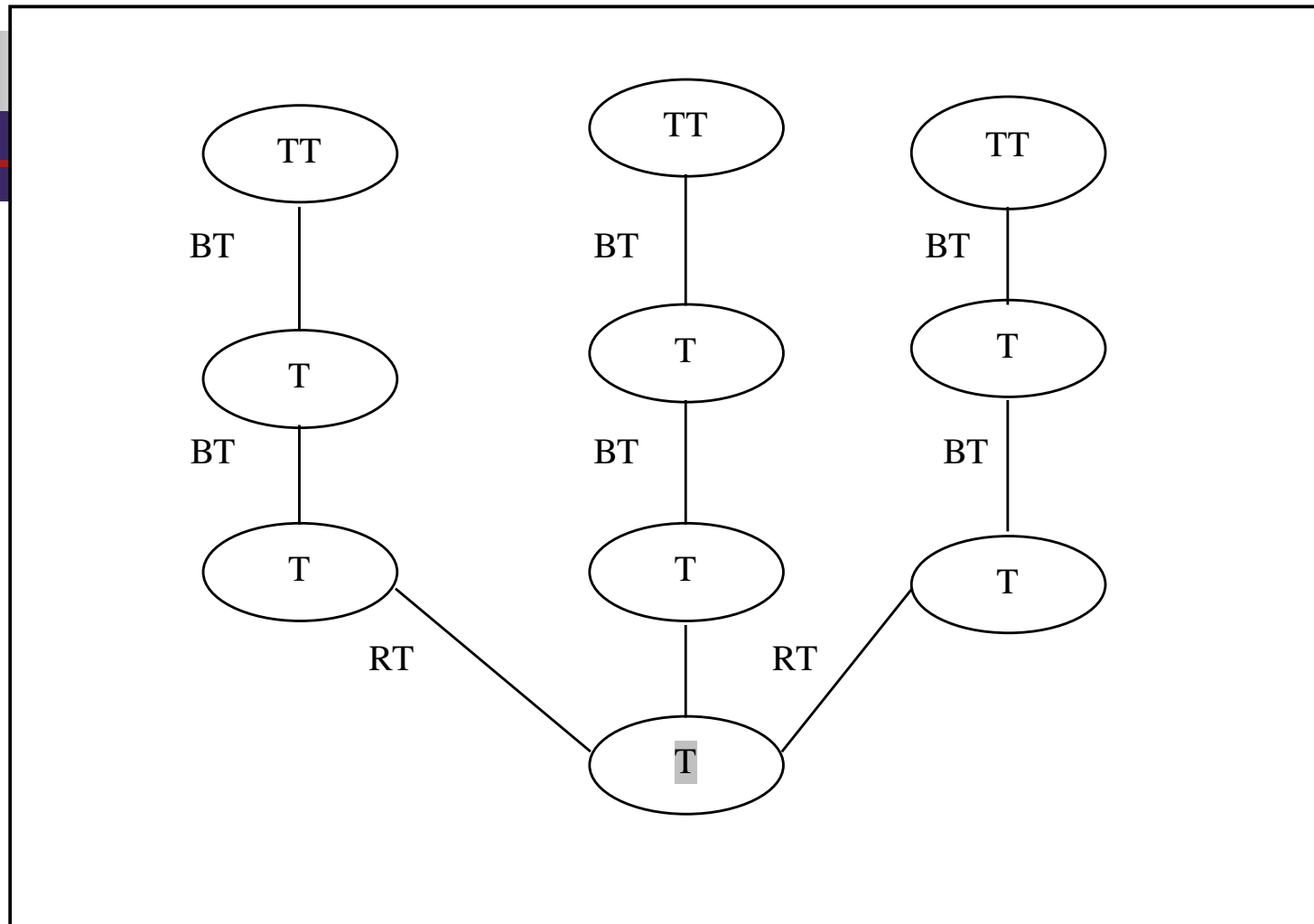
Seconda fase: revisione a partire dall'alto della macrostruttura classificatoria, con **metodo deduttivo**.



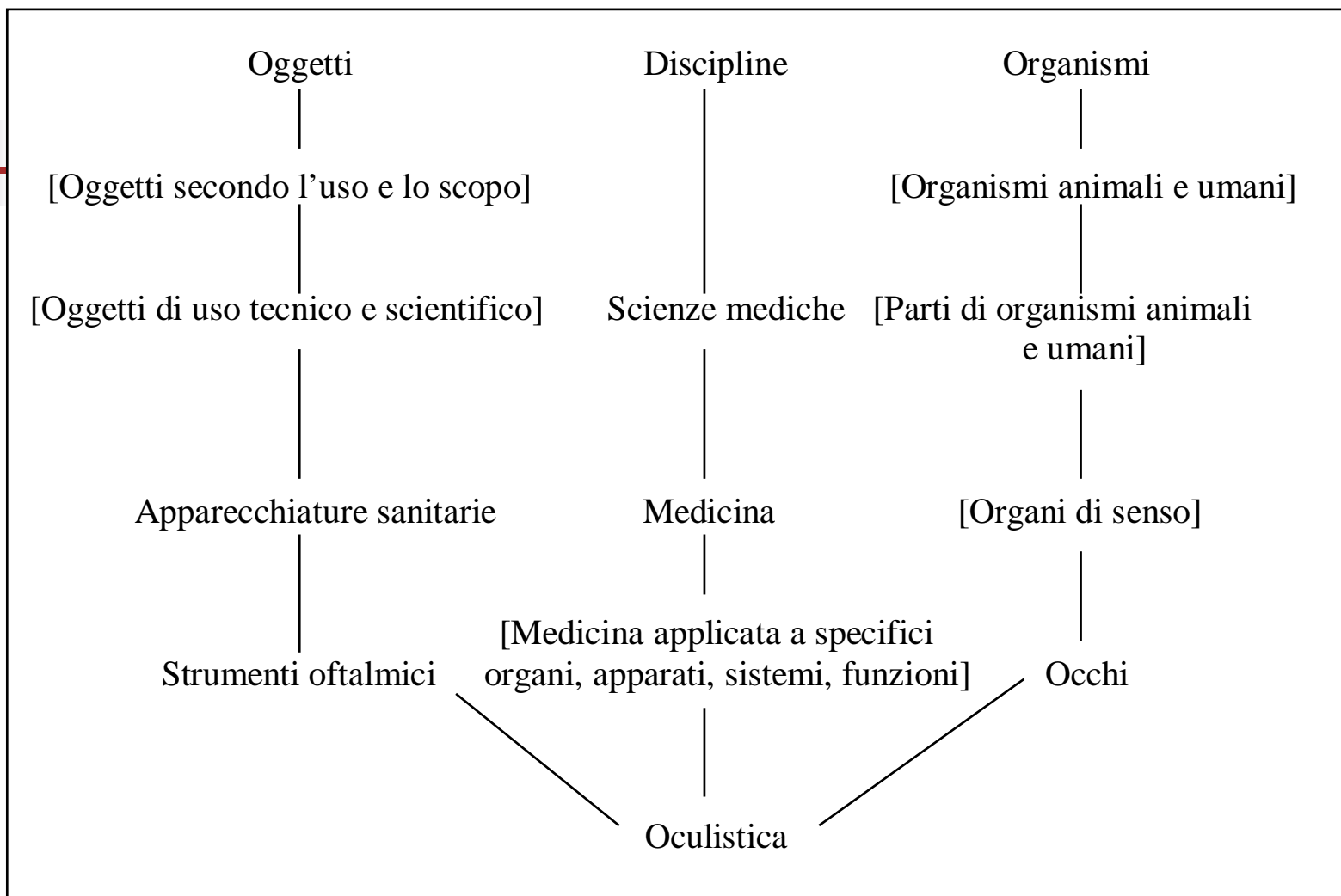
Modello di rete semantica

Il **modello di rete** è costituito:

1. dal **termine di partenza**;
2. dai **suoi termini sovraordinati**, fino al top term (relazioni gerarchiche BT/NT);
3. dai **suoi termini associati** (relazioni associative RT);
4. per **ciascun termine associato**, dai **suoi termini sovraordinati**, fino al top term (relazioni gerarchiche).



Modello di rete semantica



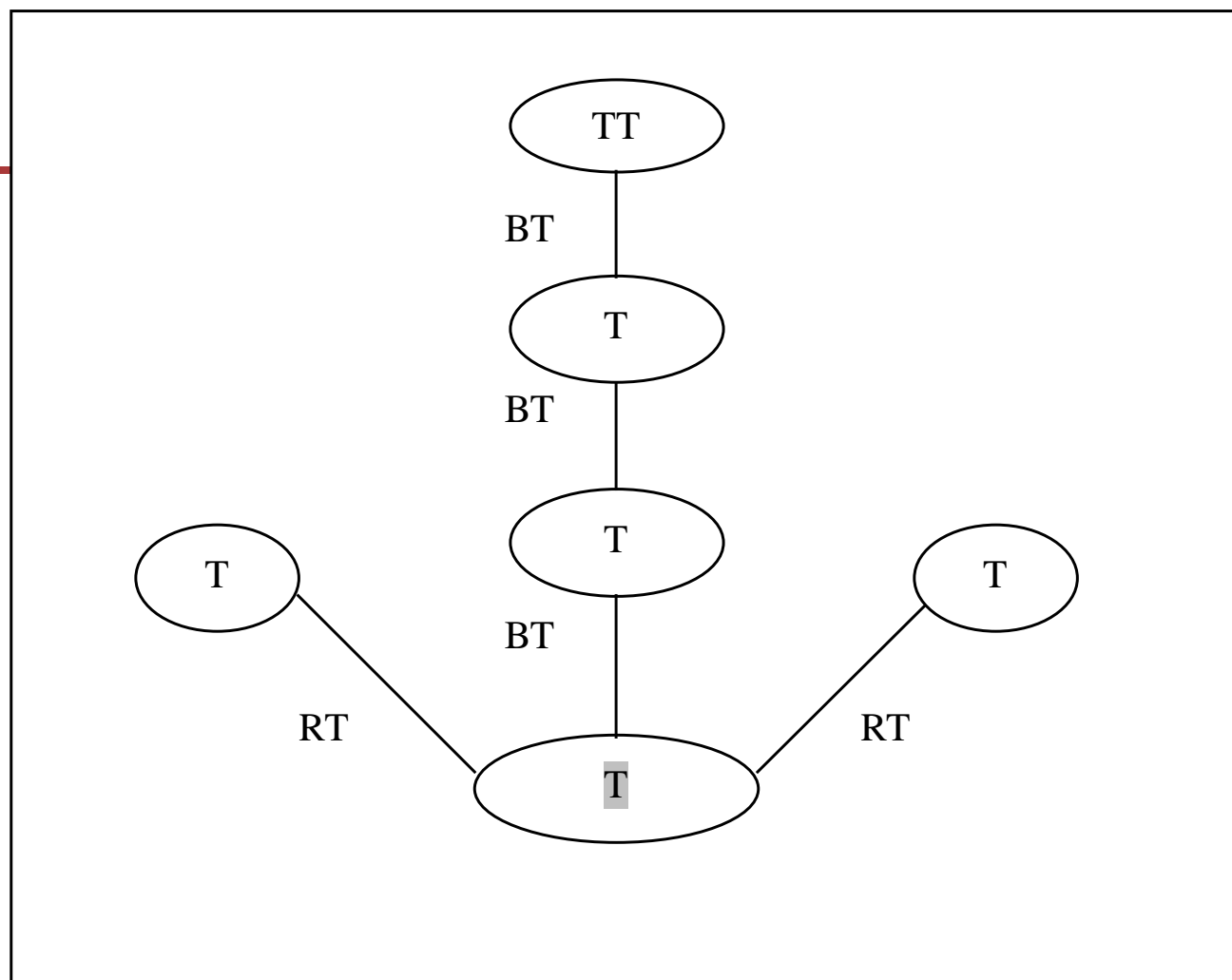
Esempio di rete semantica di un termine



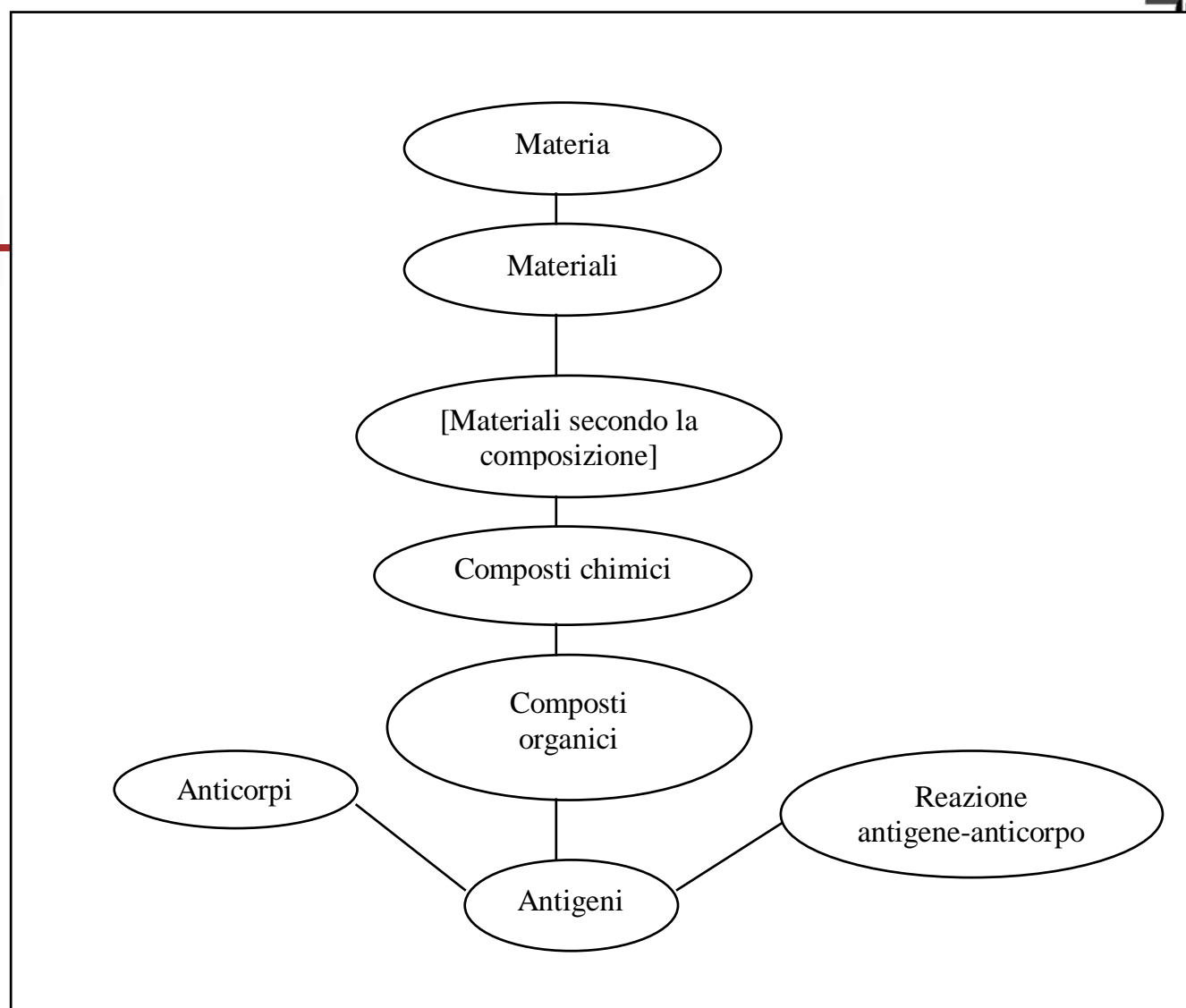
Modello di rete ad áncora

Il modello di rete semantica non prevede **termini orfani** (termini non legati al proprio termine sovraordinato).

Ma nella fase iniziale di costruzione del Prototipo è stato adottato un **modello semplificato**, più economico, denominato **modello ad ancora**.



Modello di rete “ad áncora”



Esempio di rete semantica "ad áncora"

Profilo delle categorie e delle faccette

Categorie	Faccette
Agenti	Organismi Organizzazioni Persone e gruppi
Azioni	Attività Discipline Processi
Cose	Forme Materia Oggetti Spazio Strumenti Strutture
Tempo	



Analisi a faccette

L'**analisi a faccette** è l'operazione di analizzare e classificare i termini per faccette.

La **faccetta** è un principio di divisione di una categoria o di una classe, ossia una caratteristica che guida l'attività di raggruppare e separare (classificare) i concetti.

L'analisi a faccette, che opera in modo deduttivo, dal generale al particolare, è lo strumento attraverso cui si completa, si razionalizza e si rende coerente la costruzione dal basso.



Analisi a faccette

Tre regole fondamentali:

1. la suddivisione di una classe deve avvenire applicando una sola caratteristica alla volta
2. è necessario stabilire quale sia l'ordine di precedenza
3. i termini che appartengono alla stessa sottofaccetta sono mutualmente esclusivi



Etichette di nodo

Le **etichette di nodo** sono termini o locuzioni la cui funzione è quella di **indicare una caratteristica di divisione**. Sono a tutti gli effetti termini di gerarchia che hanno sia BT che NT, ma non possono essere usati come termini di indicizzazione.

L'**elaborazione delle etichette di nodo** è stata operata a partire dai livelli più alti della gerarchia, concentrando l'attenzione sul primo livello e via via scendendo di grado, a seconda della necessità.



Etichette di nodo e loro elaborazione

Organizzazioni

- *[Organizzazioni secondo la forma]
- *[Organizzazioni di particolari categorie di persone]
- *[Parti di organizzazioni]
- *[Organizzazioni illegali]
- *Organizzazioni temporanee
- *Organizzazioni istituzionali
- *[Organizzazioni in base al settore di attività]

Etichette di nodo e loro elaborazione

*[Organizzazioni in base al settore di attività]

- **Organizzazioni religiose
- **Organizzazioni culturali
- **[Organizzazioni a carattere sociale]
- **[Organizzazioni ricreative e sportive]
- **[Organizzazioni delle arti e dello spettacolo]
- **Organizzazioni educative
- **Organizzazioni economiche
- **Organizzazioni militari
- **[Organizzazioni a carattere sanitario]
- **[Organizzazioni scientifiche]



Poligerarchia temperata

Inizialmente costruzione delle gerarchie condotta in modo **molto dettagliato, indipendentemente dalla garanzia letteraria.**

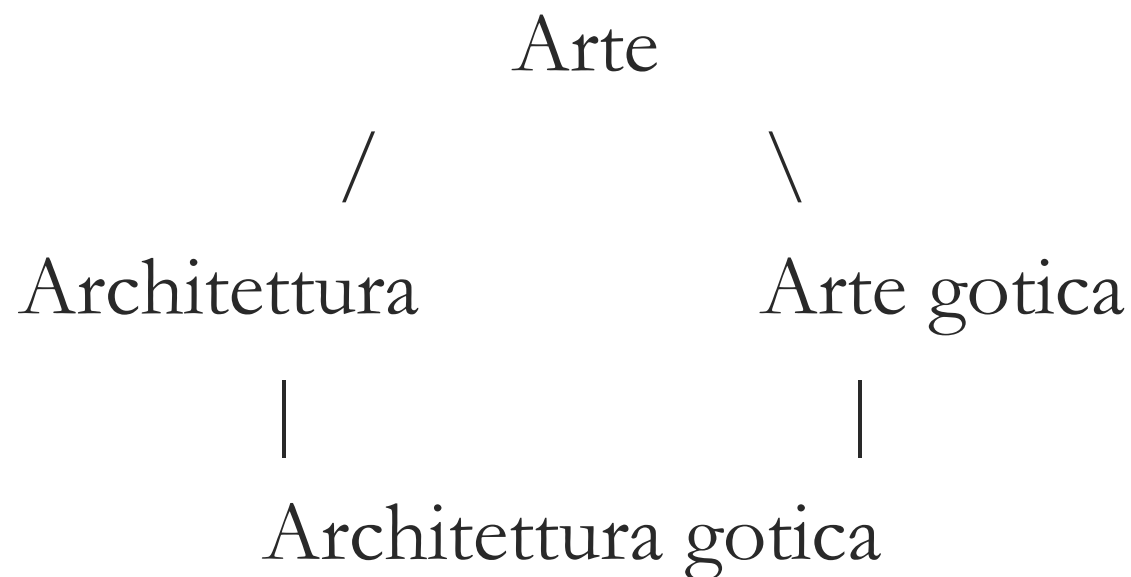
Conseguenza: **proliferazione delle gerarchie**, problema divenuto evidente con il processo deduttivo di revisione delle gerarchie dall'alto.

Soluzione imperniata su due punti cardine:

- **ricorso alla poligerarchia temperata** (ossia applicata con moderazione, entro determinati limiti e con particolari cautele
- **maggior attenzione alla garanzia letteraria.**



Conseguenze della poligerarchia temperata sulle gerarchie preesistenti



Conseguenze della poligerarchia temperata sulle gerarchie preesistenti

