



Building Event Threads out of Multiple News Articles



Xavier TANNIER and Véronique MORICEAU
LIMSI-CNRS, Univ. Paris-Sud, Orsay, France, firstname.surname@limsi.fr

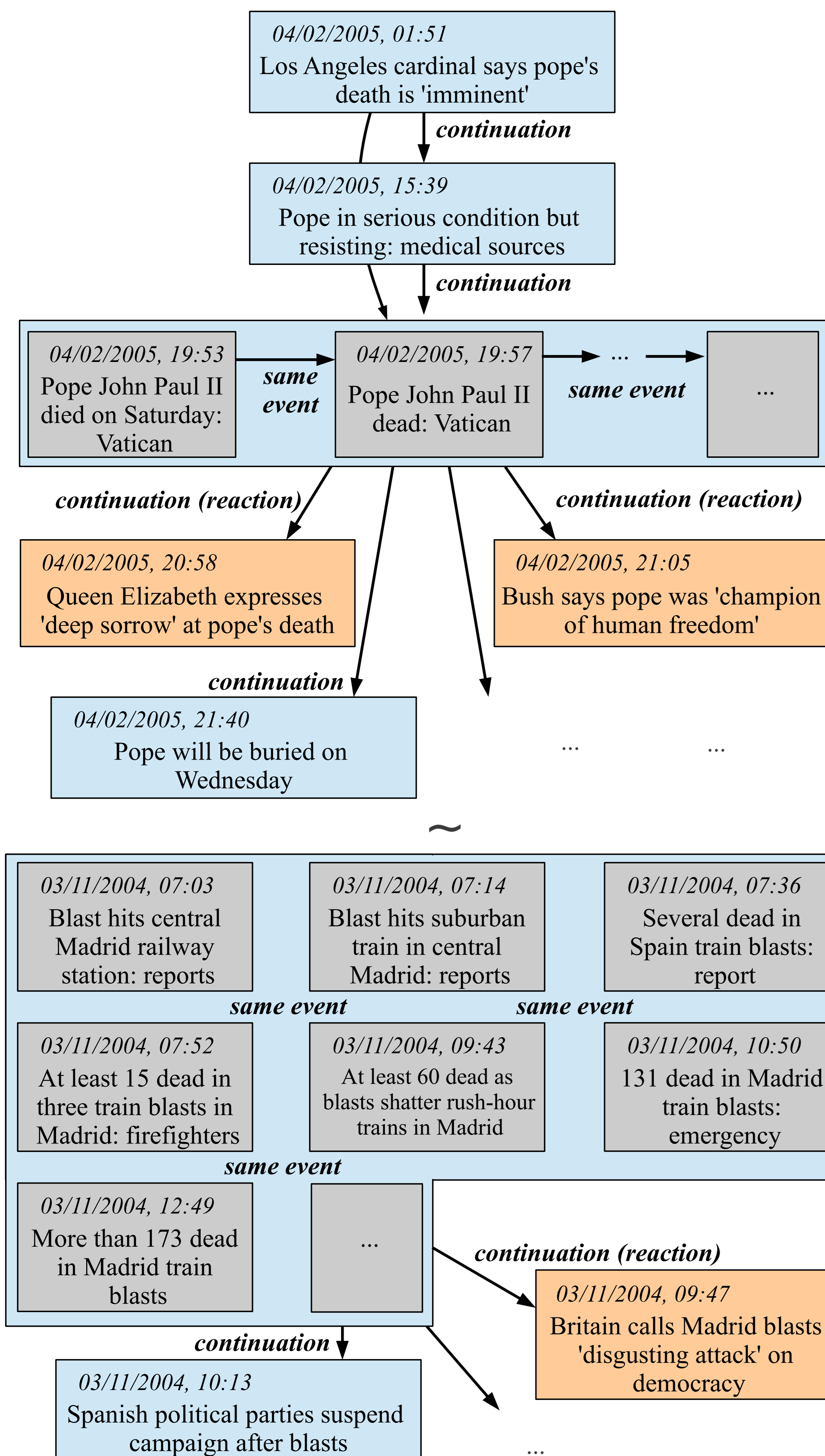
Task: Building Temporal Graphs of Articles

Vertices = news articles

Edges = relations between two documents:

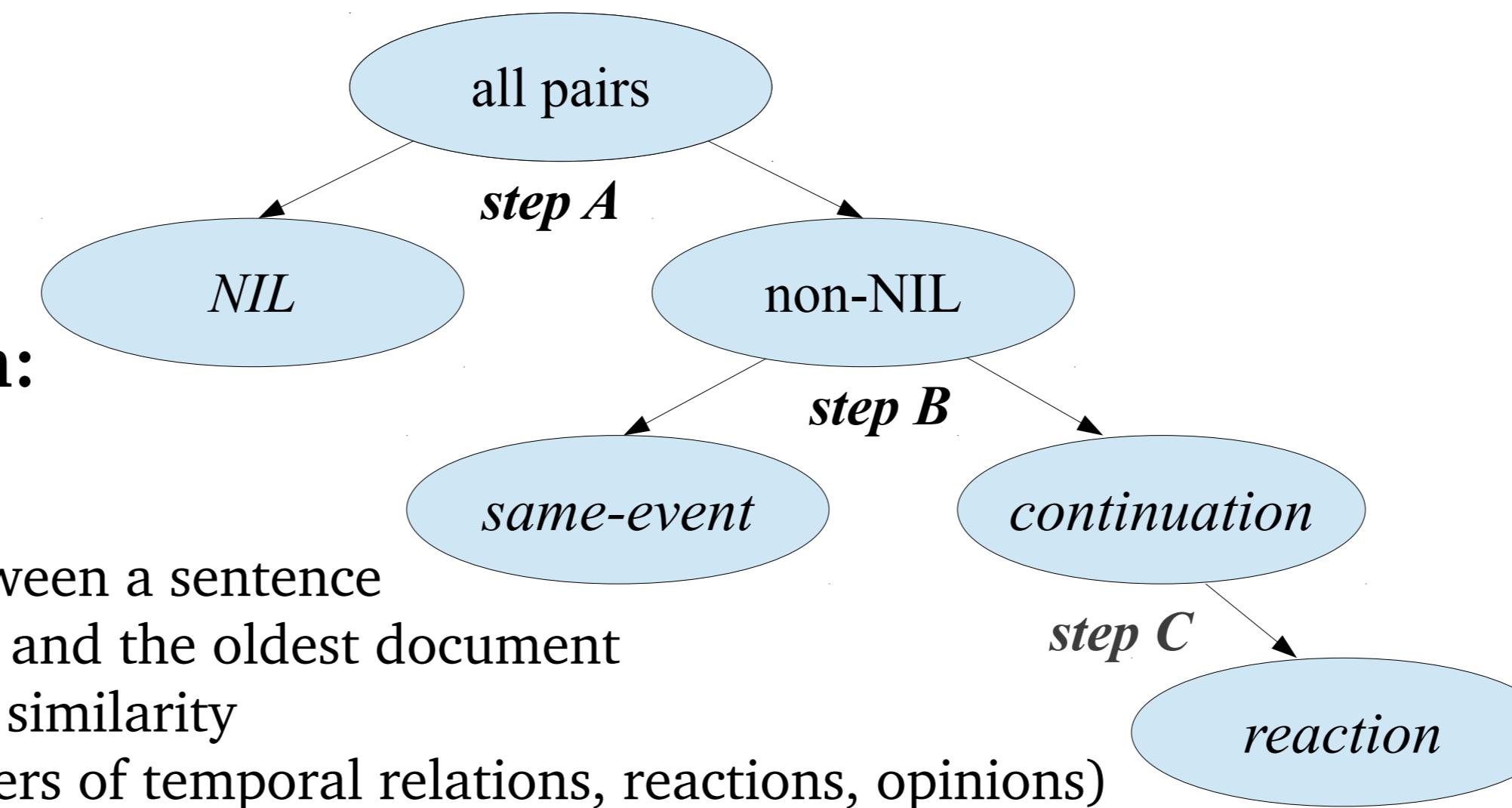
- same event
- continuation
- ↪ reaction

Objective: Proof of Concept



A cascade of classifiers

Mix of simple ML classifiers and heuristics.



Content information:

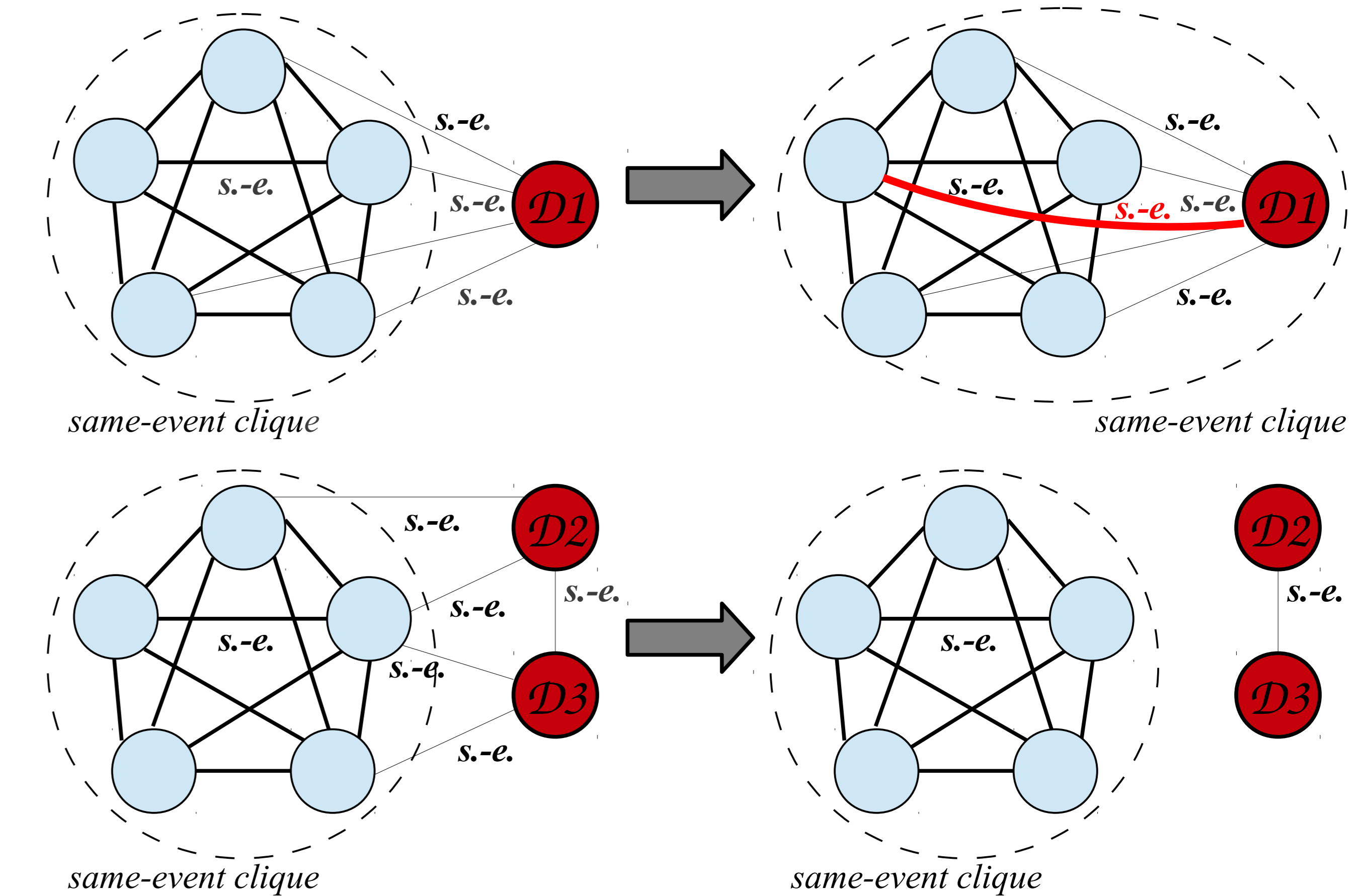
- Titles similarity
- Full text similarity
- Highest similarity between a sentence of earliest document and the oldest document
- Temporal information similarity
- Lexical features (triggers of temporal relations, reactions, opinions)

Iterative information:

- Decisions taken by previous classifiers (be more optimistic toward non-nil relations for documents having already non-nil relations)

Redundancy information:

- Build same-event cliques
- Transitive closure by vote for creating new relations or removing isolated ones
- See figure at the right



Vote for same-event transitive closure. At the top (a.), four nodes from the 5-node clique are linked to document D1, which is enough to add D1 to the clique. At the bottom (b.), only two nodes from the clique are linked to documents D2 and D3, which is not enough to add them into the clique. All edges from the clique to D2 and D3 are then deleted.

Annotation & Evaluation

Building the resources

- Manual annotation with a purposely-designed interface
- Two annotators
- Two rounds of annotation (see article for more details)
- Inter-annotator agreement Kappa = 0.83 over 150 pairs

Relation	Number of pairs	
	Learning	Test
NIL	614	304
same-event	458	304
continuation	748	386
reaction	123	59

Evaluation

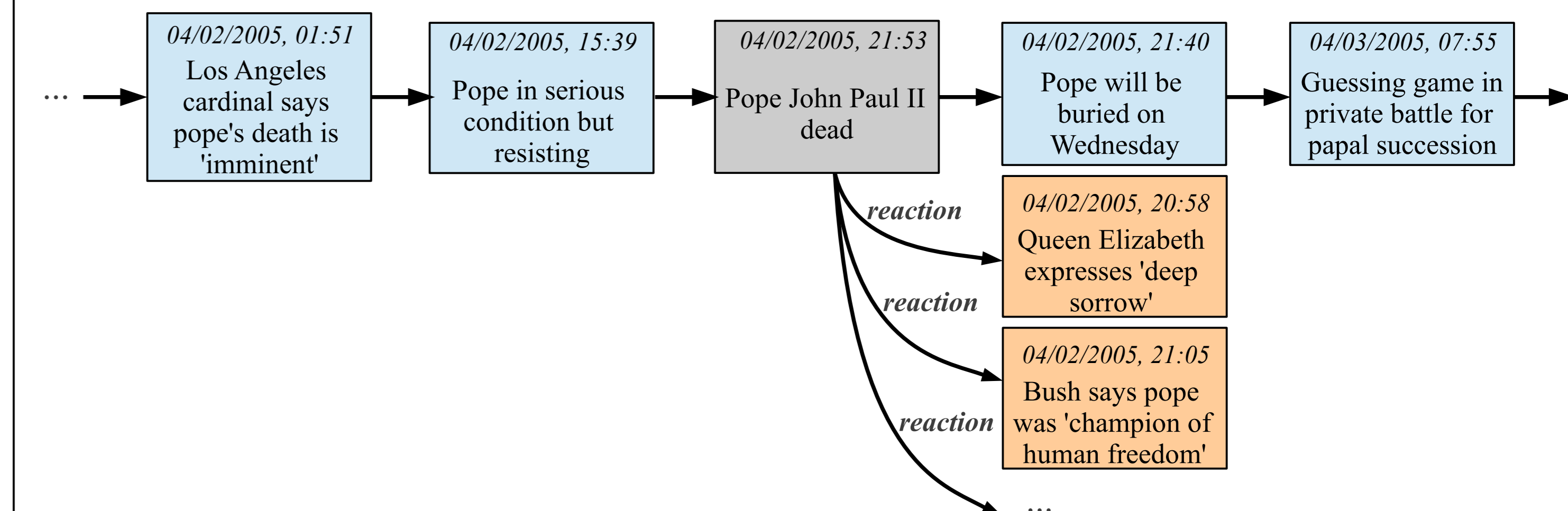
(See evaluation of each step of the process in the article)

Relation	Precision	Recall	F1
NIL	0.890	0.831	0.860
same-event	0.943	0.819	0.877
continuation (inc. reaction)	0.798	0.911	0.851
↪ reaction	0.778	0.359	0.491

Application: Event Threads

Visualization of stories and navigation through event threads

1. Start with a given article.
2. Get temporal graph with articles 7 days before and after.
3. From same-event cliques, extract only the longest or most recent article, but present date of the older article.
4. Present nodes in chronological order, isolate reactions.



- The user can visualize and navigate through this thread
- Such a temporal thread is potentially infinite. If the user navigates through the end of the 7-day window, the system must be run again on the next time span.