Granularity in Natural Language Discourse

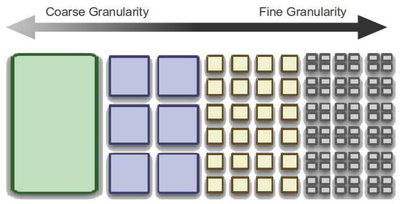
Rutu Mulkar-Mehta, Jerry R. Hobbs, Eduard Hovy

me@rutumulkar.com, hobbs@isi.edu, hovy@isi.edu

Information Sciences Institute, University of Southern California

What is Granularity?

Granularity: the level of detail of description of an event or object.



Domain	Domain elements in decreasing granularity			
Software	Services	Modules	Packages	Classes
Football	Season	Game	Drive	Play
Shopping	Shopping	Travel to Store	Walk	Move a foot
Cleaning	Clean a house	Clean a Room	Dust a cabinet	Dust a shelf

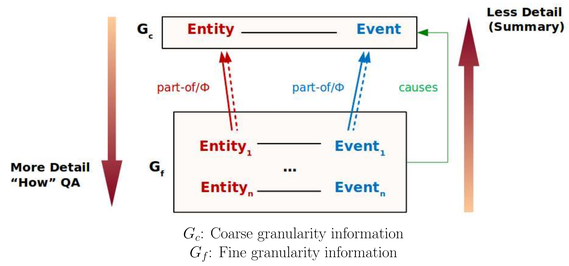
Related Work: [4], [6], [5], [1]

There are 2 types of causality:

- Sequential Causality: *The building collapsed because of the earthquake*
- Granular Causality: *The building collapsed because the roof caved in*

This poster focuses on granular causality, and how identification of granularity structure in text can help infer causal relations.

Theory of Causal Granularity



G_c: Coarse granularity information
G_f: Fine granularity information

Relevant categories of Part-of Relations [8] and Causal Relations [3] as they occur in natural language discourse.

Part-of Relations [8]	
Category	**Example**
Component-Integral	*pedal - bike*
Member-Collection	*ship - fleet*
Portion-Mass	*slice - pie*
Stuff-Object	*steel - car*
Feature-Activity	*pay - shop*
Place-Area	*LA - USA*

Causal Relations [3]		
Category	**Type**	**Example**
Causal Connectives	Prepositional	*because of, thanks to, due to*
	Adverbial	*for this reason, the result that*
	Clause links	*because, since, for*
Causation Verbs		*kill, melt*
		poison, hang
Conditionals		*If S1 then S2.*

Annotation Experiments

Articles: 37 articles from three domains (Travel, Timebank [7], Football games)
Annotators: 5 People from Mechanical Turk
Annotation Guidelines:

1. Is one paragraph a subevent of the other paragraph?
2. Did one paragraph cause the other paragraph?
3. Is one paragraph less detailed and the other paragraph more detailed?
4. Did one paragraph happen after the other paragraph?

Example Task:
1: Officials in California are warning residents of dangerous and unpredictable landslides.
2: Experts say the ground is so saturated it cannot absorb any more water. So that means soil will fall off in chunks and destroy anything in its path.

Agreement:
para2 is a subevent of para1, para2 causes para1, para2 is more detailed than para1, para2 happens before or at the same time as para1.

Verbatim comments from Annotators:
.."Without the rain or loose soil, landslide warnings don't happen"..
.."Para2 will result in Para1"..

Agreement Evaluation

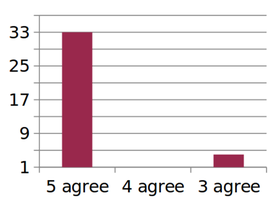
Annotation Agreement: Two annotators were considered to be in agreement if they agreed with questions 1, 3 and 4.

"Causality" was a confusing feature

1: I wanted to visit the confluence point located in the extreme southwest of Hunan Province.
2: To get to the confluence, I caught the Hong Kong-to-Shanghai intercity train on Friday afternoon.

Does 1 cause 2 or 2 cause 1?

All the annotators agreed that a sub-event explains how an event happens, or a sub-event causes an event. We counted this in lieu of our causality question (2)



Result and Analysis

Annotator ID	A2-K9	A2-CA	A2-GH	A3-CN	A3-FM	Average
A2-K9		0.82	0.76	1	0.76	**0.833314**
A2-CA	0.82		0.93	0.82	0.93	**0.876835**
A2-GH	0.76	0.93		0.76	1	**0.861683**
A3-CN	1	0.82	0.76		0.76	**0.833314**
A3-FM	0.76	0.93	1	0.76		**0.861683**

The average pairwise Kappa [2]: 0.85

Example case of disagreement
1: Some 1,500 ethnic Albanians marched Sunday in downtown Istanbul, burning Serbian flags.
2: The police barred the crowd from reaching the consulate, but allowed them to demonstrate nearby.
Positive Granularity Shift: "demonstrations" are a part of a "march"
Negative Granularity Shift: "demonstrations" happened after the "march"

Conclusions and Future Work

We are currently developing a system for automatic recognition of granularity shifts. We will compare its performance with state of the art techniques for answering causality-style questions and empirically evaluate the significance of granularity structures for automated Question Answering.

References

[1] Thomas Bittner and Barry Smith. Granular partitions and vagueness. *Proceedings of the international conference on Formal Ontology in Information Systems - FOIS '01*, pages 309–320, 2001.

[2] J Cohen. A coefficientof agreement for nominal scales. *Educational and Psychological Measurement*, 20:37–46, 1960.

[3] Roxana Girju and Dan Moldovan. Mining Answers for Causation. *Proceedings of American Association of Artificial Intelligence*, pages 15–25, 2002.

[4] Jerry R. Hobbs. Granularity. *In Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 432–435, 1985.

[5] C Maria Keet. *A Formal Theory of Granularity*. PhD thesis, Faculty of Computer Science, Free University of Bozen-Balzano, Italy, 2008.

[6] Inderjeet Mani. A Theory of Granularity and its Application to Problems of Polysemy and Underspecification of Meaning. *In Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference (KR'98)*, pages 245–255, 1998.

[7] Feng Pan, Rutu Mulkar, and Jerry R. Hobbs. Extending TimeML with typical durations of events. *Proceedings of the Workshop on Annotating and Reasoning about Time and Events - ARTE '06*, pages 38–45, 2006.

[8] Morton E. Winston, Roger Chaffin, and Douglas Herrmann. A Taxonomy of Part-Whole Relations. *Cognitive Science*, 11(4):417–444, October 1987.