Granularity in Natural Language Discourse

Rutu Mulkar-Mehta, Jerry Hobbs and Eduard Hovy
University of Southern California, Information Sciences Institute
me@rutumulkar.com, hobbs@isi.edu, hovy@isi.edu

Abstract

This paper discusses the phenomenon of granularity in natural language[1]. By 'granularity' we mean the level of detail of description of an event or object. Humans can seamlessly shift their granularity perspective while reading or understanding a text. To emulate this mechanism, we describe a set of features that identify the levels of granularity in text, and empirically verify this feature set using a human annotation study for granularity identification. This theory is the foundation for any system that can learn the (global) behavior of event descriptions from (local) behavior descriptions. This is the first research initiative, to our knowledge, for identifying granularity shifts in natural language descriptions.

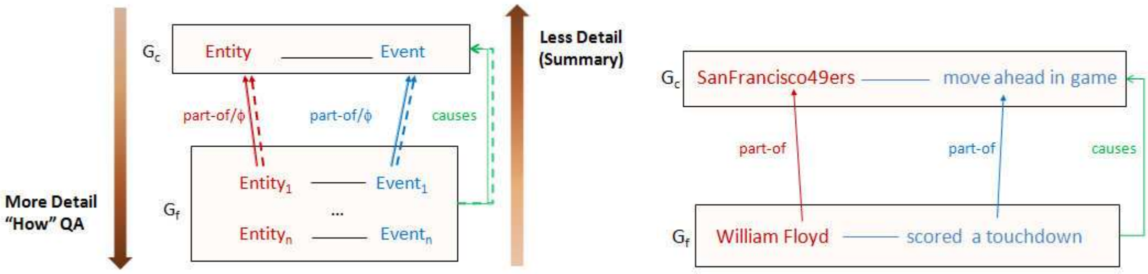
1 Introduction

Granularity is the concept of breaking down an event into smaller parts or granules such that each individual granule plays a part in the higher level event. For example, the activity of driving to the grocery store involves some fine-grained events like opening the car door, starting the engine, planning the route, and driving to the destination. Each of these may in turn be decomposed further into finer levels of granularity. For instance, planning the route might involve entering an address into GPS and following directions. The phenomenon of granularity is observed in various domains, including scientific literature, game reports, and political descriptions. In scientific literature, the process of photosynthesis on closer examination is made up of smaller individual fine-grained processes such as the light dependent reaction and the light independent reaction.

Granularity is not a new concept. It has been studied actively in various disciplines. In philosophy, Bittner and Smith (2001) have worked on formalizing granularity and part-hood relations. In information retrieval, Lau et al. (2009) have used granularity concepts to extract relevant detail of information resulting from a given search query. In theoretical computer science and ontology development, Keet (2008) has worked on formalizing the concept of entity granularity and hierarchy and applied it biological sciences. In natural language processing, Mani (1998) has worked on applying concepts of granularity to polysemy and Hobbs (1985) has worked on using granularity for decomposing complex theories into simple theories.

Although all of the above work emphasizes the importance of granularity relations for language understanding and formalization, none of it has attempted to observe whether granularity structures exist in natural language texts, explored whether granularity structures can be identified and extracted automatically, or tried to analyze how harvesting granularity relations can possibly help with other NLP problems. This paper focuses on two items: First, we present a model of granularity as it exists in natural language (Section 2); and second, we present an annotation study which we conducted to verify the proposed model of granularity in natural language (Section 3).

[1]This research was supported by the Defense Advanced Research Projects Agency (DARPA) Machine Reading Program under Air Force Research Laboratory (AFRL) prime contract no. FA8750-09-C-0172. Any opinions, findings, and conclusion or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the view of the DARPA, AFRL, ONR, or the US government.



(a) (b)

Figure 1: 1(a): Granularity in Natural Language Descriptions; 1(b): Instantiating Natural Language to the Granularity model

2 Modeling Granularity in Natural Language Texts

Humans can easily shift through various levels of granularity in understanding text. However, for automated granularity identification and extraction, it is necessary to explicitly recognize the identifiers that indicate a shift in granularity. Figure 1(a) illustrates our theory of granularity. A granularity structure exists only if at least two levels of information are present in text, such that the events at the coarse granularity can be decomposed into the events at the fine granularity, and the events at the fine granularity combine together to form at least one segment of the event at the coarse granularity. In Figure 1(a), G_c represents the phrase or sentence with coarse granularity information and G_f represents a phrase or sentence with fine granularity information. Three types of relations can exist between the objects at coarse and fine granularity: *part-whole relationships between entities*, *part-whole relationships between events*, and *causal relationships between the fine and coarse granularities*. These relations signal a shift in granularity. Instantiating text phrases into this model will expose granularities of text. For example, consider the following sentence:

The San Francisco 49ers moved ahead 7–3 11 minutes into the game when William Floyd scored a two-yard touchdown run.

The event of the player scoring a touchdown (the second clause of the sentence) is a decomposition of the event of the team moving forward in the game (the first clause), and thus a finer granularity representation of the San Francisco 49ers moving ahead in the game. When instantiated in our model of granularity (Figure 1(a)), the graphical representation is shown in Figure 1(b).

Having described the overall model of granularity, we now elaborate on the components of the granularity model, namely *part-whole relations* and *causal relations*.

2.1 Part-Whole Relations

Two types of part-whole relations are present: *meronymic* and *mereologic*. Mereology (for more details read Keet (2008)) is a partial ordering relation that is reflexive, transitive, and antisymmetric. According to the concept of mereology, if *x*, *y* and *z* are three entities, then: *x* is a part of *x*; if *x* is part of *y* and *y* is part of *z* then *x* is part of *z*; and if *x* is part of *y* then *y* cannot be part of *x*. However, various types of part-whole relations that occur in natural language, such as *member of*, do not satisfy the transitivity relation, in which case they will be mereologic but not meronymic: they might be ontologically accurate but not linguistically correct. For instance, *if John's arm is part of John, and John is a member of a football team*, the transitivity relation that *John's arm is part of a football team*, is not a valid meronymic relation. Another instance which is mereologic but not meronymic is the following: *A cup is made of steel, and steel is made of molecules. Therefore a cup is made of molecules.* The concept of mereology does not

reflect the way *part of* is used in natural language, and so mereology cannot be used for linguistic based research.
One of the early works on part-whole relations in natural language (meronymy) Winston et al. (1987) was later refined in their empirical experiments Chaffin et al. (1988). Winston et al. discuss meronymic relations and a taxonomy for representing them. They introduce six types of part-whole relationships: (i) Component-Integral (e.g., *pedal* is a component of the integral *bike*), (ii) Member-Collection (e.g., a *ship* is a member of the collection, a *fleet*), (ii) Portion-Mass (e.g., a *slice* is a portion of the mass, a *pie*), (iv) Stuff-Object (e.g., *steel* is one of the ingredients/stuff of the object *car*), (v) Feature-Activity (e.g., *paying* is one of the features of the whole activity of *shopping*), (vi) Place-Area (e.g., *Everglades* is a place within the area of *Florida*). The definition and classification in Winston et al. (1987) for part-whole relations is very relevant for language based analysis of part-whole relations. For granularity identification in our work, the Feature-Activity type relation is used as the part-whole relation for events, and the rest are part-whole relations for entities.

2.2 Causal Relations

Girju and Moldovan (2002) provide a broad compilation of causality research ranging from philosophy, planning in AI, commonsense reasoning, and computational linguistics. Causation in computational linguistics is the only form of causality that is relevant for granularity identification and extraction. The following are the categories of causal constructs relevant for granularity identification and extraction:

- Causal Connectives: These are usually prepositional (such as *because of, thanks to, due to*), adverbial (such as *for this reason, the result that*), or clause links (such as *because, since, for*).
- Causation Verbs: These usually have a causal relation integrated with the verb. For example, *kill*, *melt* (represent a causal link with the resulting situation), *poison*, *hang*, *clean* (represent a causal link with the a part of the causing event)
- Conditionals: Girju and Moldovan (2002) describe conditionals as complex linguistic structures typically of the form *If S1 then S2*. These structures represent causation, temporal relations, among other relations, and are very complex structures in language.

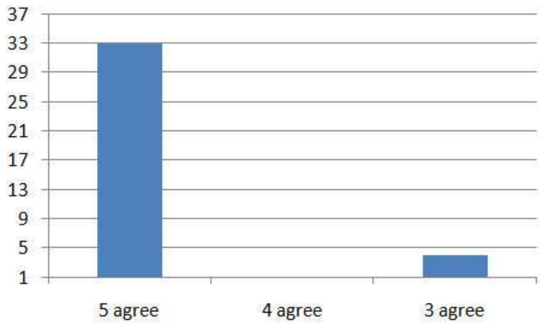
3 Evaluation of the Granularity Model in Natural Language

We conducted an evaluation study to judge the "goodness" of the granularity model proposed. In this study the annotators were asked to annotate granularity relations between two given paragraphs. Paragraph-based analysis was preferred to event-word-based analysis because people reason much more easily with paragraph descriptions than with individual event mentions [2]. The annotation set consisted of paragraph pairs from three domains: travel articles (confluence.org), Timebank annotated data Pan et al. (2006), and Wikipedia articles on games. We selected a total of 37 articles: 10 articles about travel, 10 about games, and 17 from Timebank. Both paragraphs of a given question were selected from the same article and referred to the same overall concept.

3.1 Annotation Task

The articles were uploaded to Mechanical Turk and were annotated by non-expert annotators (regular Turkers). The entire set of 37 articles was annotated by 5 people. The annotators were given a pair of paragraphs and were asked four questions about the relations between them: (i) Is one paragraph a subevent of the other paragraph?, (ii) Did one paragraph cause the other paragraph?, (iii) Is one paragraph less detailed and the other paragraph more detailed?, (iv) Did one paragraph happen after the other paragraph? They were then presented with the comments of other annotators, and asked whether they agreed

[2]This was deduced as a result of an earlier annotation study for granularity identification using individual words as events.



(a)

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

(b)

Figure 2: 2(a) shows the Inter-Annotator agreement for 37 articles and 2(b) shows the Pairwise Kappa Agreement for 37 articles and 5 annotators

with any of the other annotations or explanations. The annotators were asked to provide a justification of their choices.

3.2 Results

The Kappa statistic (Cohen (1960)) is the standard for measuring inter-annotator agreement: $k = \frac{(p(a)-p(e))}{(1-p(e))}$, where *p(a)* is the observed agreement and *p(e)* is the chance agreement between annotators. More refined than simple Percentage Agreement, Kappa corrects for chance agreements.

In our study, two annotators were considered to be in agreement if they agreed with questions (i) Subevents, (iii) More or less detail and (iv) Sequence. Unfortunately question (ii) Causality, as provided to the annotators, could not be taken into account for agreement measurement as individuals had different conceptualizations of causality, and a crisp definition of causality was not provided to them. For instance, consider the following two paragraphs:

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.*
Analysis: Some annotators annotated *para2 causes para1*, providing the explanation that the goal para1 could be achieved due to the events of para2. Others annotated *para1 causes para2*, providing the justification that the events of para2 only exist to fulfill the original goal para1. We are interested in the first type of causality, i.e., causality which explains *how* a given event happens. 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 (ii).

Figure 2(a) shows the overall agreement of the five annotators on the 37 articles and Figure 2(b) shows the pairwise Kappa agreement for the five annotators. All the annotators agreed in 33/37 cases (23 article pairs were annotated as having a granularity shift, 10 articles were annotated as having no granularity shift). The average pairwise Kappa was 0.85. If the newspaper articles were removed, the overall agreement was 100% for all the annotators. High agreement implied good quality of the annotation guidelines, and provided evidence that people shift through various levels of granularity while reading and understanding text.

3.3 Analysis of the Causes of Disagreement

Where disagreements occurred, different interpretations of the same text were observed to be a major cause. All these disagreements were limited to the newspaper articles. For instance, consider the following:

1: Some 1,500 ethnic Albanians marched Sunday in downtown Istanbul, burning Serbian flags.
2: The police barred the crowd from reaching the Yugoslavian consulate in downtown Istanbul, but allowed them to demonstrate on nearby streets.
Positive Granularity Shift: Some annotators commented that "demonstrations" happen as a part of a "march". So, para2 is a sub-event of para1.
Negative Granularity Shift: Other annotators felt that para2 happened after para1, and so there was no granularity shift.

Overall, we can observe that although disagreement arises due to individual and unique interpretations of text, people agree based on the discriminating features provided to them (part-whole relations and causality) when identifying granularity shifts. This shows that part-whole relations and causality provide a good set of features for identifying granularity shifts.

4 Conclusion and Future Work

In this paper we present the phenomenon of granularity as it occurs in natural language texts. We validate our model of granularity with the help of an annotation study. We are currently developing a system for automatic granularity extraction. We will compare its performance with state of the art techniques for answering causality-style questions to empirically evaluate the significance of granularity structures for automated Question Answering.

References

Bittner, T. and B. Smith (2001). Granular partitions and vagueness. *Proceedings of the international conference on Formal Ontology in Information Systems - FOIS '01*, 309–320.

Chaffin, R., D. J. Herrmann, and M. E. Winston (1988). An empirical taxonomy of part-whole relations: Effects of part-whole relation type on relation identification. *Language and Cognitive Processes 3*(1).

Cohen, J. (1960). A coefficientof agreement for nominal scales. *Educational and Psychological Measurement 20*, 37–46.

Girju, R. and D. Moldovan (2002). Mining Answers for Causation. *Proceedings of American Association of Artificial Intelligence*, 15–25.

Hobbs, J. R. (1985). Granularity. *In Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, 432–435.

Keet, C. M. (2008). *A Formal Theory of Granularity*. Ph. D. thesis, Faculty of Computer Science, Free University of Bozen-Balzano, Italy.

Lau, R. Y. K., C. C. L. Lai, and Y. Li (2009). Mining Fuzzy Ontology for a Web-Based Granular Information Retrieval System. *Lecture Notes in Computer Science*, 239–246.

Mani, I. (1998). 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)*, 245–255.

Pan, F., R. Mulkar, and J. R. Hobbs (2006). An Annotated Corpus of Typical Durations of Events. *In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)*, 77–83.

Winston, M. E., R. Chaffin, and D. Herrmann (1987, October). A Taxonomy of Part-Whole Relations. *Cognitive Science 11*(4), 417–444.