

7. Übungsblatt - Abgabe: 20.6.2014, 16:00

Answers can be written in either English or German (English is preferred).

THE TEXT

The following text is taken from the Penn Discourse Treebank. Some entities are highlighted and marked with a grammatical function (Subject, Object, X=other). These are the ones to include in the Entity Grid.

1. (**Researchers-S**) at American Telephone & Telegraph Co. 's (**Bell Laboratories-X**) reported (**they-S**) raised the electrical (**current-carrying capacity-O**) of new superconductor (**crystals-X**) by a factor of 100, moving the materials closer to commercial use.
2. The (**scientists-S**) said (**they-S**) created small changes in the crystal-lattice structures of the (**superconductors-X**) to raise the amount of current that single (**crystals-S**) could carry to 600,000 amps per square centimeter in a moderately strong magnetic field.
3. The (**scientists-S**) said (**they-S**) made the advance with yttrium-containing (**superconductors-O**) cooled to liquid-nitrogen temperature, or minus 321 degrees Fahrenheit.
4. Their (**report-S**) appears in today's issue of the journal Nature.
5. (**Superconductors-S**) conduct electricity without resistance when cooled.
6. A family of ceramic (**superconductors-S**) discovered during the past three years promise new technologies such as cheaper electrical generation – but only if their (**current-carrying capacity-O**) can be raised.
7. The AT&T (**advance-S**) shows how one aspect of the (**current-carrying problem-O**) can be overcome.
8. But “(**it-S**) won't lead to imminent use” of new (**superconductors-O**), cautioned (**Robert B. van Dover-S**), one of the AT&T (**researchers-X**).
9. The reason: the (**current-carrying capacity-S**) of (**multi-crystal materials-X**) of (**superconductors-X**) remains too low for most practical uses because of so-called weak links between (**crystals-X**).
10. Such (**multi-crystal materials-O**) will probably be needed for commercial applications.

Exercise 1: Entity Grid Model

(a) For the text above, produce an entity grid, as described in lecture and as shown in Barzilay & Lapata 2008.

(b) Calculate the probabilities of the following transitions for this text: (S S), (O -), (X - -). Note:

- Use only the entities highlighted in the text - your grid should have 10 entities.
- Assume that *scientists* and *researchers* are the same entity.
- Assume that coreference has been handled (i.e. *they* can be replaced with the entity it refers to).
- If an entity appears twice in the same sentence, only mark it once. If the GFs are not the same, use the highest ranking.

Exercise 2: Discourse Relations

In the PDTB, four different types of discourse relation between two arguments are annotated.

- Explicit Relations
- Implicit Relations
- Alternate Lexicalizations (AltLex)
- Entity Relations (EntRel)

From the text above, find one example of each relation type. Write down the relation type, the connective, and the two arguments. For Implicit relations, fill in the connective you think should be there. For AltLex, indicate what part of the text signals the relation. For EntRel, indicate which entity the two arguments have in common.

Hint: In most cases, the two arguments will either be within one sentence or across two adjacent sentences.