

Annotation Guidelines for Entity Link Tracking (LNK) Version 3.0 20040401

1 Intro

The Entity Link Tracking task is designed to group all references to an entity and its properties together.

Basic Concepts

- An **Entity** is an object or set of objects in the world. Entities can be referenced in a text by their name, a nominal phrase, or a pronoun. We will distinguish between nominal mentions that are descriptions and those that ascribe properties to an entity.
- **Coreference** is the connection established between all referential mentions of an entity.
- A **Link** is the connection established between an entity and an attributive mention that describes it or ascribes a property to it.
- A **Composite Entity** results from linking an entity to all attributive mentions of its properties.
- Every entity is assigned an **Entity Class** that describes the kind of reference it makes to a thing in the world.
- Every Composite Entity will be assigned a **Composite Entity Class** which corresponds to the entity class of its referential entity.

2 Links: Entity Grouping

The Linking task involves grouping all references to a single entity and all its properties together into a Composite Entity. This is a coreference task for the entire text.

2.1 Attributive Mention and Entity Link

We will Link all Specific and Generic entities with the predicates and other attributive (ATR) mentions that ascribe properties to them. This will insure that each Composite Entity is made up of all strings of text used to refer to or describe a given entity in the discourse. The following constructions should be examined for possible links.

2.1.1 Predicate Complements

When a property is ascribed to an entity in an affirmatively asserted predicate complement, we will link the ATR mention with the entity to which that property is ascribed.

[The resort] is [a popular summer and early winter skiing destination].

Modal and Negative Constructions

Sentences with modal constructions and those that negatively assert properties will not motivate link tags. For example, we would not link “John” and “a real linguist” in the following sentences.

John is not a real linguist.

John might be a real linguist.

John can be a real linguist.

John should be a real linguist.

2.1.2 Apposition

In the EDT task, apposition is tagged in a formulaic manner where the entire extent and first element are Specific references to an entity and the second element is an Attributive mention. We will link these attributive mentions to the entities they describe.

[Norbert Karlsboeck], [mayor of Kaprun]

2.1.3 Premodifiers

All premodifiers are tagged as Attributive. We will link premodifiers with referential entities when the premodifier is ascribing a property derived from the entity to the thing it modifies. For example,

[Russia]’s Federal Security Service arrested Pope on suspicion of trying to illegally buy plans for a secret high-speed torpedo used by the [Russian] navy.

The specific, referential entity Russia and the attributive mention Russian will be linked as Russian is ascribing “Russia” qualities to “the navy”.

2.1.4 Quotes

Properties are often ascribed to entities within quoted text. In these cases, we will establish links between the entities and those attributes when they conform to the above rules for taggability.

Oleg Plotnikov, the state prosecutor, said that [Pope] was [a career intelligence officer who was attempting to purchase the manufacturing technology]

3 Cross-type Metonymy

Cross-Type Metonymy occurs when a single Composite Entity is composed of entities with different EDT types. The most common example occur between ORGs and the FACs they occupy. One way to think about this is to emphasize the agentive nature of the organization and the artifactual quality of the facility. At the EDT stage, these two characteristics of the entity will be tagged separately as type ORG for textual references referring to the organization and as type FAC when the mention refers to the physical building. We will group entities of different types together into a Composite Entity by creating links between them when they refer to different aspects of the same underlying object in the real world.

I'm reporting from the steps of [the Supreme Court]....

...[The Supreme Court] decided not to allow a recount in Florida.

In this example, the first mention of 'Supreme Court' is of EDT type FAC. The second mention is of EDT type ORG. Each of these mentions evokes a different aspect of the 'Supreme Court' entity and should be linked to reflect this connection.

4 Composite Entity Class Assignment

Unlinked mentions with ATR class must be assigned a Composite Entity Class. The CE Class should correspond to the Entity Class of the thing it is referencing in the real world that is not explicitly mentioned in the text.

In the following examples, it should be assumed that the larger context does not contain any referential (SPC, GEN, USP) mentions of the ATR entity.

[US] troops

[car] manufacturer

[Microsoft] employees

[UN] member countries