Annotation Guidelines for

Relation Detection and Characterization (RDC)

Version 4.3.2 - 20040401

1 Introduction

The goal of RDC is to detect and characterize relations of the targeted types between EDT entities. Every relation takes two primary arguments (the two entities that it links) and must be assigned one of the seven syntactic relation class types. For all of these types, with the exception of Verbal, the relation extent is limited to the noun phrase. Subtypes will be assigned to every relation further characterizing the identified relationships. For each Type, there is a set of possible Subtypes.

Basic Concepts

As a general rule, annotators should select the relation type and subtype that conveys the most information.

> *a US marine*

> [a [US] marine]

In this example, the annotator must choose between a citizen relationship and an employment relationship between the 'marine' and the 'US'. The correct relationship is the employment as it expresses more information.

2 Types and Subtypes

A. Physical (PHYS)

Physical relations describe physical proximities of taggable entities.

1 Located

The Located relation captures the exact location of an entity. However, if an entity is located in a geographical region like a lake, a river, or a mountain, it should be reported as a Located relation even if the text does not explicitly refer to the shores of the lake, the banks of the river, or the foothills of the mountain.
> *a military base in Germany*
> [PHYS.Located("a military base in Germany", "Germany")]
>
> *St. Vartan's Cathedral, on Second Avenue*
> [PHYS.Located("St. Vartan's Cathedral, on Second Avenue", "Second Avenue")]
>
> *a station at the top of the mountain*
> [PHYS.Located("a station at the top of the mountain", "the top of the mountain")]

2 Near

Near indicates that an entity is explicitly near another entity, but not actually in that location or part of that location.

a town some 50 miles south of Salzburg
[PHYS.Near("a <u>town</u> some 50 miles south of Salzburg", "<u>Salzburg</u>")]

3 Part-whole

Part-Whole characterizes physical relationships between entities and their parts.

Norwalk, Conn.
[PHYS.Part-Whole("<u>Norwalk</u>, Conn.", "<u>Conn</u>.")]

the top of the mountain
[PHYS.Part-Whole("the <u>top</u> of the mountain", "the <u>mountain</u>")]

a state within the former Soviet Union
[PHYS.Part-Whole("a <u>state</u> within the former Soviet Union", "the former <u>Soviet Union</u>")]

B. Personal/Social (PER-SOC)

Personal/Social relations describe the relationship between entities of type PER. No other entity type is allowed as an argument of these relations. The order of the arguments does not impact relations of this type. We record only that there exists a relationship between the entities.

1 Business

Business captures the connection between two entities in any professional relationship. This includes boss-employee, lawyer-client, co-workers, political relationships, etc.

their colleagues
[PER-SOC.Business("<u>their</u>", "their <u>colleagues</u>")]

his lawyer
[PER-SOC.Business("<u>his</u>", "his <u>lawyer</u>")]

a spokesman for the senator
[PER-SOC.Business("a <u>spokesman</u> for the senator", "the <u>senator</u>")]

2 Family

Family captures the relation between an entity and another entity with which it is in any familial position.

relatives of the dead
[PER-SOC.Family("<u>relatives</u> of the dead", "the <u>dead</u>")]

his wife
[PER-SOC.Family("his <u>wife</u>" "<u>his</u>")]

his ailing father
[PER-SOC.Family("his ailing <u>father</u>", "<u>his</u>")]

3 Other

Other is reserved for all Social relationships that do not cleanly fit into the subtypes above.

> *his roommates*
> [PER-SOC.Other("his", "his roommates")]

C. Employment/Membership/Subsidiary – (EMP-ORG)

This relation includes

- **Employment** captures the relationship between PERs and the ORG or GPE by which they are employed. It is important to note that this relation is only valid when the PER is paid by the ORG. (PER – ORG), (PER – GPE)

- **Subsidiary** captures the ownership, administrative, and other hierarchical relationships between organizations and between organizations and GPEs. (ORG – ORG), (ORG – GPE)

- **Membership** captures the relationship between an agent (PER, ORG, GPE) and an organization of which they are a member. (ORG – ORG), (GPE – ORG), (PER – ORG)

1 Employ-exec(s)

This subtype describes relations between persons and organizations where the person holds a managerial position such as CEO, president, vice-president, director, leader, or head.

For example:

> *George Bush, the US president,*
> [EMP-ORG.Employ-exec ("the US president,", "US")]

> *the CEO of Microsoft*
> [EMP-ORG.Employ-exec("the CEO of Microsoft", "Microsoft")]

2 Employ-staff

This subtype is for relationships between organizations and GPEs and persons who fill general staff positions within them.

> *Mr. Smith, a senior programmer at Microsoft...*
> [EMP-ORG.Employ-staff("a senior programmer at Microsoft", "Microsoft")]

3 Employ-undetermined

At times the context does not give you enough information to determine whether an individual is performing a managerial or general staff position within an organization. Employ-undetermined is for these relations.

> *Microsoft spokesman, Bob Jones*
> [EMP-ORG.Employ-undetermined("Microsoft spokesman", "Microsoft")]

4 Member-of-group

Member relations include organization membership such as political party membership, church membership, and so on. For example:

> *an activist for Peace Now*
> [EMP-ORG.Member-of-group("an <u>activist</u> for Peace Now", "<u>Peace Now</u>")]

Organizations and GPEs can be members of other Organizations. The Member Subtype describes the relationship between these entities. For example:

> *three permanent UN member countries, the US, England, and China,*
> [EMP-ORG.Member-of-group("*three permanent UN member <u>countries</u>*", "<u>UN</u>")]

5 Subsidiary

Subsidiary characterizes the relationship between a company and its parent company.

> *Shares of Disney, parent company of ABC, are up five eighths.*
> [EMP-ORG.Subsidiary ("<u>ABC</u>", "parent <u>company</u> of ABC")]

It also describes the relationship between a department of an organization and that organization. This includes the organizational aspect of GPEs. For example:

> *New York police*
> [EMP-ORG.Subsidiary("New York <u>police</u>", "<u>New York</u>")]

> *Microsoft's accounting department*
> [EMP-ORG.Subsidiary("Microsoft's accounting <u>department</u>", "<u>Microsoft</u>")]

> *The U.S. Congress decided to veto the ecology bill.*
> [EMP-ORG.Subsidiary (The U.S. <u>Congress</u>, <u>U.S</u>)]

6 Partner

Partner characterizes the collaborative relationship between two agents (PER, ORG, GPE).
> *Microsoft and NBC reached an agreement to establish a partnership.*
> [EMP-ORG.Partner("<u>Microsoft</u>","<u>NBC</u>")]

7 Other

Other is reserved for relationships between PER, ORG, and GPE that do not fit into the other schema.

D. Agent-Artifact (ART)

Agent-Artifact describes the relationship between agentive entities and artifacts.

1 User/Owner

An agent is in a Possessor/Owner relationship with an artifact when that agent is the owner of the artifact or has possession of or habitually uses it. In the following example, it is not explicitly clear whether I own or rent the house. Possessor/Owner can be applied to either relationship.

My house is in West Philadelphia.
[ART.User-Owner("my","my house")]

2 Inventor/Manufacturer

An agent is in an Inventor/Manufacturer relationship with an artifact when that agent caused the artifact to come into being. For example,

Rubin Military design, the makers of the Kursk
ART.Inventor-Manufacturer("the makers of the Kursk", "the Kursk")]

US helicopters flew over northern Iraq.
ART.Inventor-Manufacturer("US", "US helicopters")]

British Airways bought seven Beoing 777s.
ART.Inventor-Manufacturers("Beoing","Beoing 777s")]

3 Other

Other is reserved for any Agent-Artifact relations that do not fall under the other two subtypes.

E. PER/ORG Affiliation (Other-AFF)

PER/Org Affiliation describes relationships between entities that are not captured by other relation types.

1 Ethnic

Ethnic describes the relationship between Person(s) and the collective PER group to which they are identified.

Cuban-American people
[OTHER-AFF.Ethnic("Cuban-American people", "Cuban-American")]

Arab people
[OTHER-AFF.Ethnic("Arab people", "Arab")]

Please note that many ethnic designations are expressed between a person and the GPE of which they are citizens or residents. For these, we establish GPE Affiliation relations with the subtype Citizen/Resident.

American school children
[GPE-Affiliation.Citizen/Resident("American school children", "American")]

2 Ideology

Ideology describes the relationship between Person(s) and the collective PER/ORG group(s) defined by coherent ideological systems to which they are identified by themselves or the article.

Christian people

[OTHER-AFF.Ideology("Christian people", "Christian")]

3 Other

Other should be used for all PER/ORG Affiliation relations that do not fit cleanly into any other categories. Many of the relations that fall under this subtype will be cases where a PER or ORG modifies another entity. The intended meaning of this construction is often unclear. This subtype can also be filled with relations that have type overlap. For example:

Gore supporters
[OTHER-AFF.Other("Gore supporters", "Gore")]

Jewish people
[OTHER-AFF.Other("Jewish people", "Jewish")]

Please note that employment relations should be tagged as employment relations.

Microsoft programmer
[EMP-ORG.Employ-staff("Microsoft programmer", "Microsoft")]

This example should be tagged as an Employ-staff. This follows our general rule of tagging for the most information.

F. GPE Affiliation (GPE-AFF)

GPE Affiliation describes the relationship between entities of type PER and ORG and GPEs when more than one aspect of the GPE is referenced by the context of the text.

1 Citizen/Resident

Citizen/Resident describes the relation between a PER and the GPE in which they have citizenship or in which they live.

U.S. businessman Edmond Pope
[GPE-AFF.Citizen("U.S. businessman", "U.S")]

2 Based-in

Organizations are not always located in the GPE in which they are based. We distinguish between the physical locations of an ORG with their GPE of origin with the Based-In Subtype.

Salzburg Red Cross officials
[GPE-AFF.Based-In("Salzburg Red Cross officials", "Salzburg")]

Please note that entities of type PER can never be an argument of a GPE.Affiliation.Based-in relation.

3 Other

Other should be used for all GPE Affiliation relations that do not fit cleanly into any other categories. Many of the relations that fall under this subtype will be

cases where a GPE modifies another entity. The intended meaning of this construction is often unclear.

> *More and more US companies are moving their operations to Mexico.*
> [GPE-AFF.Other("US <u>companies</u>", "<u>US</u>")]

Note that the relationship between GPEs and government organizations should be EMP-ORG.Subsidiary.

G. Discourse (DISC)

A Discourse (DISC) relation is one where a semantic part-whole or membership relation is established only for the purposes of the discourse. The whole or group referred to is not an official entity relevant to world knowledge. Instead, it has been constructed for solely the purposes of discursive efficiency.

> *Many of these people*
> [DISC("<u>Many</u> of these people", "these <u>people</u>")]

> *each of whom*
> [DISC("<u>each</u> of whom", "<u>whom</u>")]

> *each of our parishes*
> [DISC("<u>each</u> of our parishes", "our <u>parishes</u>")]

> *the dead included dozens of Austrians and Germans, as well as Japanese tourists and American soldiers from a military base in nearby Germany*
> [DISC("<u>dozens</u> of Austrians and Germans", "the <u>dead</u> included dozens of Austrians and Germans, as well as Japanese tourists and American soldiers from a military base in nearby Germany")]
> [DISC("Japanese <u>tourists</u>", "the <u>dead</u> included dozens of Austrians and Germans, as well as Japanese tourists and American soldiers from a military base in nearby Germany")]
> [DISC("American <u>soldiers</u> from a military base in nearby Germany ", "the <u>dead</u> included dozens of Austrians and Germans, as well as Japanese tourists and American soldiers from a military base in nearby Germany")]

2.2 Relation Lexical Condition

Unlike Entities and Events, Relations have no actual anchor in the text. The six relation classes are intended to provide justification for the tagging of each relation. With the exception of the Verbal class, the syntactic classes strictly limit taggable relations to those that are explicitly stated and fall within the maximum extent of the longest noun phrase.

A. Possessive

Possessive indicates the syntactic structure where the first noun or pronoun is in the possessive case.

> *America's Department of Defense.*
> [EMP-ORG.Subsidiary("America's <u>Department of Defense</u>", "<u>America</u>")]

Nathan Myhrvold, Microsoft's chief scientist.
[EMP-ORG.Staff("Microsoft's chief <u>scientist</u>", "<u>Microsoft</u>")]

B. Preposition

Preposition indicates a taggable relation between a head noun and a prepositional phrase that modifies it.

Officials in California are warning residents.
[PHYS.Located("<u>officials</u> in California", "<u>California</u>")]

The CEO of Microsoft
[EMP-ORG.Employ-exec("The <u>CEO</u> of Microsoft", "<u>Microsoft</u>")]

C. PreMod

PreMod relations are those motivated by the construction of a proper adjective or proper noun premodifier and the head noun it modifies.

The American envoy left the talks early.
[EMP-ORG.Employ-staff("The American <u>envoy</u>", "<u>American</u>")]

Palestinian leaders
[EMP-ORG.Employ-exec("Palestinian <u>leaders</u>", "<u>Palestinian</u>")]

New York police
[EMP-ORG.Subsidiary("New York <u>police</u>", "<u>New York</u>")]

C.1 Nested Premodifier Relations

The nesting of premodifiers determines which mentions should be included as arguments of taggable relations.

US President George Bush
[[[<u>US</u>] <u>President</u>] <u>George Bush</u>]

There are three EDT mentions in this string of text. Since 'US' is modifying the 'President' mention, we will mark the extent of this EMP-ORG.Employ-executive relation as

US President
[EMP-ORG.Employ-executive("US <u>President</u>", "<u>US</u>")]

D. Formulaic

For these standard constructions, we will capture the following relations.

Reporter sign-off
Jane Clayson, ABC News, South Lake Tahoe.
[PHYS.Located("<u>Jane Clayson</u>", "South <u>Lake Tahoe</u>")]
[EMP-ORG-Employ-staff("<u>Jane Clayson</u>", "<u>ABC News</u>")]

Addresses

Mary Smith, Medford, Mass. I feel we should…
[PHYS.Located("<u>Smith</u>", "<u>Medford</u>, Mass.")]
[PHYS.Located("<u>Medford</u>, Mass", "<u>Mass</u>")]

Elected officials

Senate Majority Leader Trent Lott (R-Miss.)
[EMP-ORG.Member-of-Group("Senate Majority <u>Leader Trent Lott</u>", "<u>R</u>")]
[GPE-AFF.Citizen-resident("Senate Majority Leader <u>Trent Lott</u>", "<u>Miss</u>")]

E. Verbal

Verbal relations are those motivated by a taggable relation between entities that appear within the extent of a construction headed by a verb and are not contained within a single noun phrase. The full extent of the relation will be the entire phrase.

Death Valley is in the Mojave Desert.

Credit Suisse is in a coalition of banks against money laundering.

Coca Cola Co. is based in Atlanta.

the crowd trapped inside the compartment…
[PHYS.Located("the <u>crowd</u> trapped inside the compartment", "the <u>compartment</u>")]

159 people, including 8 Americans
[DISC("8 <u>Americans</u>", "159 <u>people</u>, including 8 Americans")]

the private-sector body based in Norwalk, Conn
[GPE-AFF.Based-In("the private-sector <u>body</u> based in Norwalk, Conn", "<u>Norwalk</u>, Conn")]