ERGO-Annotation: Annotation for Noun Phrases Contained in Eligibility Criteria

Samson Tu,[1] Mor Peleg,[2] Simona Carini,[3] Daniel Rubin,[1] Ida Sim[3]

[1]Center for Biomedical Informatics Research, Stanford University, Stanford, CA

[2]Department of Management Information Systems, University of Haifa, 31905, Israel

[3]Division of General Internal Medicine, UCSF, San Francisco, CA

I.   Introduction

This document presents a simple summary of ERGO-Annotation – a method of using noun phrases contained in eligibility criteria to annotate the criteria.

II.   Objectives

ERGO-annotation captures the noun phrases contained in eligibility criteria in order to have some idea about what the criteria "are about". With the noun expressions, we can query across criteria to find studies whose target populations have certain characteristics. With the noun phrases, we can make use of terminological reasoning to aggregate studies (e.g., all studies whose target populations have cardiovascular disease).

III.   Using ERGO-annotation with ERGO

We attach Simple_Statement_Annotation to a Simple Clinical Statement, a Comparison_Statement_Annotation to a Comparison Statement, and a set of (Simple_Statement_Annotation or Comparison_Statement_Annotation) to a Compound Statement. We do not annotate Semantically_Connected_Statement or Partially_Specified_Statement.

IV.   The Structure of ERGO Annotations

There are two types of annotations: Simple_Statement_Annotation and Comparison_Statement_Annotation.

A.   *Simple_Statement_Annotation*

Simple_Statement_Annotation refer to an augmented_noun_phrase (i.e., a Noun_Phrase or Context-Dependent_Noun_Phrase).

B.   *Comparison_Statement_Annotation*

A Comparison_Statement_Annotation is an annotation that says "some aspect" of noun phrase is in a quantitative relationship to a quantity. Comparison_Statement_Annotation compares a Noun_Phrase or Context-Dependent_Noun_Phrase to a quantity (part of ERGO Expressions) using a comparator symbol.
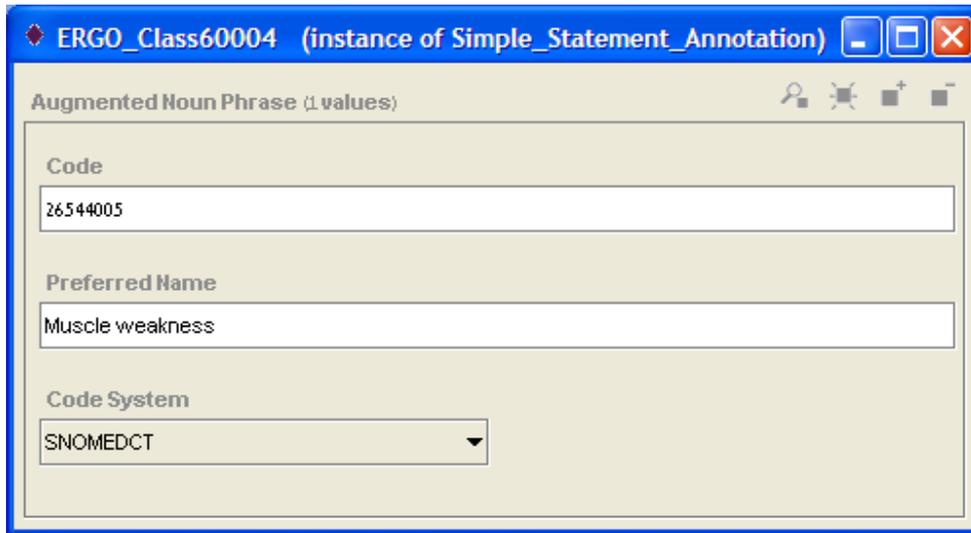
C.   *Noun Phrases*

Noun_Phrases are primitive or post-coordinated terms that should be linked to standard terminologies (e.g., SNOMED-CT) for interoperability. There are three subclasses of noun phrases: Primitive noun phrases which represent terms from vocabularies, Logical combinations of noun phrases and Modified noun phrases. These are explained in the ERGO Technical Document. We will use SNOMED-CT Concept model attribute (410662002) and Unapproved attribute (408739003) subtrees as our initial set of modifiers for Modified noun phrases.

Context-Dependent_Noun_Phrases - This class is used to encode, in simple way, SNOMED finding with context such as "clinical finding absent with associated-finding noun_phrase" or "family history of noun phrase. Context-Dependent_Noun_Phrases are noun phrases modified by a context_modifier. Context modifiers are taken from SNOMED-CT's situation with explicit context

V.   Examples of annotations of eligibility criteria

a) A Simple_Statement_Annotation annotated with a primitive noun phrase: Presence of weakness

b) A Simple_Statement_Annotation annotated with a Context-Dependent_Noun_Phrase: "clinical finding absent: Frank hematuria"
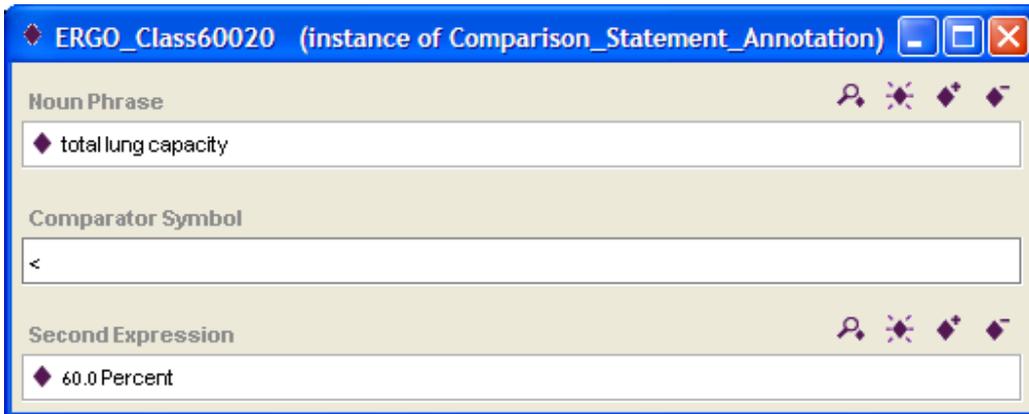

c) A comparison statement annotation for a criterion that compares the queried value of total lung capacity to a threshold

ERGO_Class60020 (instance of Comparison_Statement_Annotation)

Noun Phrase

◆ total lung capacity

Comparator Symbol

<

Second Expression

◆ 60.0 Percent

d) A Simple_Statement_Annotation annotated with a Logical_Combination_Noun_Phrase: oral anti-diabetic excluding metformin

ERGO_Class60026 (instance of Simple_Statement_Annotation)

Augmented Noun Phrase (1 values)

Modifier Term

◆ and

Create Instance

Concepts

◆ oral anti-diabetic
◆ Not metformin

Not metformin (instance of Negated_Noun_Phrase...

Modifier Term

◆ Not

Create Instance

Root Concept

◆ metformin