## Word-Sense Disambiguation

One of the most difficult problems MetaMap deals with is ambiguity, which is rampant in the UMLS. For example, *cold* is 7-ways ambiguous, meaning that each of the 7 CUIs shown below along with their preferred names includes the string *cold* (or *colds*) as a synonym:

- 1. Cold Sensation (C0234192)
- 2. Cold Temperature (C0009264)
- 3. Common Cold (C0009443)
- 4. Cold Therapy (C0010412)
- 5. Cold brand of chlorpheniramine-phenylpropanolamine (C0719425)
- 6. Colds homeopathic medication (C1949981)
- 7. Chronic Obstructive Airway Disease  $\left(\mathrm{C0024117}\right)$

The last four senses of *cold* are manually suppressed from MetaMap's datafiles, leaving only the first three.

MetaMap's Word-Sense Disambiguation (WSD) module, invoked with  $\neg y,$  can be used to eliminate such ambiguity.

For example, because of the extreme ambiguity of each word in the string

fixed dose combination treatment groups

MetaMap by default will generate 1,323 mappings; however, WSD will choose one concept for each of the five input words and limit the output to a single mapping.

Note that invoking WSD will add some overhead to processing.