Ambiguity in the UMLS Metathesaurus

2006 Edition

Alan R. Aronson and Sonya E. Shooshan

May 10, 2006

1. Introduction

The UMLS® Metathesaurus® contains a significant amount of ambiguity. For example, the string "Cold" (or "cold" or "COLD") occurs in six distinct concepts with six distinct meanings. The purpose of this report is to examine ambiguity in the 2006AA release of the Metathesaurus in the context of its effect on natural language processing (NLP) applications.

Until the 2004AC release of the UMLS Knowledge Sources, ambiguity was denoted explicitly by appending an ambiguity designator, a number in angle brackets, to the end of an ambiguous string. Thus the ambiguity for "cold" was denoted by 'Cold <1>', 'Cold <2>', 'COLD <3>', etc. Now ambiguity is computed by finding concepts with strings that differ only with respect to case.[1]

Table 1 shows that the degree of Metathesaurus ambiguity has grown over the years and was particularly explosive in 2005, partly due to the direct computation of ambiguity mentioned above.

	2001	**2002**	**2003**	**2004**	**2005**	**2006**
Strings with an ambiguity designator	12,840	13,837 (+8%)	16,438 (+19%)	21,295 (+30%)	N/A	N/A
Concepts with one or more ambiguity	9,637	10,328 (+7%)	12,397 (+20%)	16,775 (+35%)	36,133 (+115%)	44,591 (+23%)
Concepts with one or more non-suppressible ambiguity	8,074	8,754 (+8%)	10,416 (+19%)	12,387 (+19%)	33,513 (+171%)	40,977 (+22%)
Cases of ambiguity	5,571	6,014 (+8%)	7,204 (+20%)	10,018 (+39%)	22,218 (+122%)	27,599 (+24%)
Cases of non-suppressible ambiguity	5,311	5,752 (+8%)	6,824 (+19%)	9,521 (+40%)	20,996 (+121%)	25,290 (+20%)

Table 1. **Measures of ambiguity in the UMLS Metathesaurus**

(Percentage changes are computed relative to the previous year.) Although the overall growth in 2006 is relatively modest, the number of cases of ambiguity of higher degree has risen alarmingly. (See Table 2 and Table 3 below.) This is largely due to the inclusion of HL7, both versions 2.5 and 3.0. For example, the highest non-suppressed degrees of ambiguity in 2006AA are for "other"

---

1. Note that AMBIGSUI.RRF or AMBIG.SUI cannot be used for this purpose because they do not conflate case.

(degree 40 of which 36 are from HL7V2.5 and 1 from HL7V3.0), "retired code" (degree 39 all of which are from HL7V3.0) and "unknown" (degree 36 of which 33 come from HL7V2.5 and 1 from HL7V3.0).

Examining the cases of ambiguity more closely, consider the *degree* of ambiguity, i.e., the number of ways a string is ambiguous or, equivalently, the number of concepts in which it (or one of its case variants) occurs.[1] For example "other" has degree 54 in 2005. Table 2 contains the distribu-

Degree of ambiguity	2003 cases	2004 cases	2005 cases	2006 cases
92				1
56	1			
54		1	1(0%)	
39				1
36				1
24				1
23	1	1(0%)	1(0%)	
20				1
19				1
18	2	2(0%)	1(-50%)	1(0%)
17	58	58(0%)	2(-97%)	
16	2	2(0%)	1(-50%)	2(+100%)
15				1
14				1
13				1
12			1	1(0%)
11				3
10				4
9			6	13(+117%)
8	3	3(0%)	10(+233%)	23(+130%)
7	6	7(+17%)	11(+57%)	28(+155%)
6	8	7(-13%)	24(+243%)	66(+175%)
5	24	22(-8%)	54(+145%)	158(+193%)
4	79	76(-4%)	208(+174%)	452(+117%)
3	594	600(+1%)	1,239(+107%)	1,868(+51%)
2	5,541	7,250(+31%)	20,659(+185%)	24,971(+21%)
1	885	1,989(+125%)		
Total	**7,204**	**10,018(+39%)**	**22,218(+122%)**	**27,599(+24%)**

Table 2. **Metathesaurus ambiguity distribution by degree**

_____

1. The computation of the degree of an ambiguity was corrected in 2002. As a result, there are some differences from previous editions of this report in the counts reported in the tables.

tion of ambiguities in the Metathesaurus according to degree. Note that an ambiguity of degree one is not actually an ambiguity. In 2004 and before, for example, 'Abbreviations <1>' is not ambiguous since there were no other 'Abbreviations <n>' strings in the Metathesaurus.

Ignoring suppressible synonyms produces the more realistic distribution shown in Table 3. Most

Degree of ambiguity	2003 cases	2004 cases	2005 cases	2006 cases
40				1
39				1
36				1
24				1
23	1	1(0%)	1(0%)	
20				1
19				1
18	1	1(0%)	1(0%)	1(0%)
17			2	
16	1	1(0%)		
15				1
13				1
12			1	1(0%)
11				1
10				4
9			5	9(+80%)
8	3	3(0%)	8(+167%)	16(+100%)
7	4	2(-50%)	5(+150%)	16(+220%)
6	1	1(0%)	7(+600%)	39(+457%)
5	11	7(-36%)	31(+343%)	123(+297%)
4	44	42(-5%)	156(+271%)	360(+131%)
3	473	416(-12%)	1,000(+140%)	1,586(+59%)
2	4,935	4,309(-13%)	19,779(+359%)	23,126(+17%)
1	1,350	4,738(+251%)		
Total	**6,824**	**9,521(+40%)**	**20,996(+121%)**	**25,290(+20%)**

Table 3. **Metathesaurus ambiguity distribution after removing suppressibles**

of the ambiguity of degree greater than eight has disappeared, and all of that would disappear if appropriate strings were marked as suppressible. For example, "other" now has degree two. One of these cases, 'Other' (C0205394), is legitimate; and the other, 'Other location of complaint' (C0220886), should be marked as suppressible. Suppressible synonyms are ignored for the remainder of this report.

Section 2 of this report describes only the most notable cases of ambiguity in the Metathesaurus, i.e., the cases of degree six or more. The bulk of the cases are now reported automatically by the Migration Assistant, a tool developed generally for annotating ambiguity and specifically for the purpose of marking appropriate cases as suppressible. Finally, Section 3 provides some conclusions of the study.

2. Higher Degree Metathesaurus Ambiguity

Ambiguous English Metathesaurus strings are described in this section in decreasing order of degree of ambiguity. Only those cases of degree six or more are covered. See Migration Assistant reports for cases of ambiguity of lesser degree.

In all cases, suppressible synonyms are ignored as is done in Table 3. Ambiguous forms for concepts shown in bold should be marked as suppressible. Recommendations for cases which are not clear are introduced with the word *consider*. Ambiguous forms for concepts shown in italics should be marked as suppressible in MetaMap only.

2.1 "other" (degree 40)

Except for 'Other', the remaining cases should be marked as suppressible because they mean something more specific than "other". The concepts involved are

1. C0205394: Other
2. **C0220886: Other location of complaint**
3. **C1521979: Other Routes of Drug Administration**
4. **C1546380: Other - Event Reason**
5. **C1546725: Other Specimen Source Code**
6. **C1546836: Other - Special Program Code**
7. **C1546840: Other - Publicity Code**
8. **C1546902: Other - Diagnosis Classification**
9. **C1546930: Other - Report Source**
10. **C1547110: Other - Modality**
11. **C1547196: Other - Organization unit type**
12. **C1547233: Other - Triage Code**
13. **C1547241: Other - Newborn Code**
14. **C1547267: Other - Risk Management Incident Code**
15. **C1547272: Other - Incident Type Code**
16. **C1547281: Other - Production Class Code**
17. **C1547292: Other - Recreational Drug Use Code**
18. **C1547304: Other - Precaution Code**
19. **C1547309: Other - Patient Condition Code**
20. **C1547994: Other - Diagnostic Service Section ID**
21. **C1549063: Other - Notify Clergy Code**
22. **C1549104: Other - Administrative Gender**
23. **C1549110: Other - Marital Status**
24. **C1550146: Other - Substance Type**
25. **C1556042: Other - Relationship**
26. **C1556043: Other - Religion**

27. **C1556044: other - No Information**
28. **C1556045: Other - What subject filter**
29. **C1556046: Other - Employment Status**
30. **C1556048: Other - Contact Role**
31. **C1556049: Other - Mail Claim Party**
32. **C1556050: Other - Living Dependency**
33. **C1556051: Other - Event Consequence**
34. **C1556052: Other - Indirect exposure mechanism**
35. **C1556053: Other - Action Taken in Response to the Event**
36. **C1556054: Other - Status of Evaluation**
37. **C1556055: Other - Causality Observations**
38. **C1556056: Other - Job Status**
39. **C1556057: Other - Immunization Registry Status**
40. **C1561608: Other - Mode of Arrival**

2.2 "retired code" (degree 39)

All thirty-nine cases should be suppressed because they are specific kinds of "retired code". Their concepts are

1. **C1550523: SUGG - Retired Code**
2. **C1551999: DEP - Retired Code**
3. **C1552818: ordered - Retired Code**
4. **C1552819: unordered - Retired Code**
5. **C1552841: tbody - Retired Code**
6. **C1552842: tfoot - Retired Code**
7. **C1552843: thead - Retired Code**
8. **C1553543: PRA - Retired Code**
9. **C1553692: PSYCHCF - Retired Code**
10. **C1553693: INVOICE - Retired Code**
11. **C1553694: TRAN - Retired Code**
12. **C1554193: REFR - Retired Code**
13. **C1554965: TBLHDR - Retired Code**
14. **C1554966: TBLDATA - Retired Code**
15. **C1554967: TBLROW - Retired Code**
16. **C1554968: TBLCOL - Retired Code**
17. **C1554969: TBLCOLGP - Retired Code**
18. **C1554970: DOCTBL - Retired Code**
19. **C1554971: DOCCNTNT - Retired Code**
20. **C1554972: DOCPARA - Retired Code**
21. **C1554973: DOCLIST - Retired Code**
22. **C1554974: DOCLSTITM - Retired Code**
23. **C1555027: RACT - Retired Code**
24. **C1555339: LINKHTML - Retired Code**
25. **C1555340: LOCALATTR - Retired Code**
26. **C1555341: LOCALMRKP - Retired Code**
27. **C1555450: RETRO - Retired Code**
28. **C1555451: PROA - Retired Code**
29. **C1555452: BONUS - Retired Code**

30. **C1555453: RECOV - Retired Code**
31. **C1555454: EDU - Retired Code**
32. **C1555477: DEDUCT - Retired Code**
33. **C1555478: COPAY- Retired Code**
34. **C1555549: DOSEIND- Retired Code**
35. **C1555550: WGHT - Retired Code**
36. **C1555551: HGHT - Retired Code**
37. **C1555698: ERR - Retired Code**
38. **C1555699: WARN - Retired Code**
39. **C1555700: INFO - Retired Code**

2.3 "unknown" (degree 36)

Except for 'Unknown' (occurs twice), the remaining cases should be marked as suppressible because they mean something more specific than "unknown". The concepts involved are

1. C0439673: Unknown
2. **C1521803: Unknown Route of Drug Administration**
3. **C1546837: Unknown - Special Program Code**
4. **C1546841: Unknown Publicity Code**
5. **C1547283: Unknown - Production Class Code**
6. **C1547294: Unknown - Recreational Drug Use Code**
7. **C1547306: Unknown - Precaution Code**
8. **C1547312: Unknown - Patient Condition Code**
9. **C1548340: Unknown - Allergy Severity**
10. **C1548502: Unknown - Vaccines administered**
11. **C1548543: Unknown - Living Will Code**
12. **C1548550: Unknown - Organ Donor Code**
13. **C1549064: Unknown - Notify Clergy Code**
14. **C1549105: Unknown - Administrative Gender**
15. **C1549115: Marital Status - Unknown**
16. **C1549625: Unknown - Ethnic Group**
17. **C1556120: Unknown - Religion**
18. **C1556121: Unknown - Event reason**
19. **C1556122: Unknown - Relationship**
20. **C1556123: Unknown - Employment Status**
21. **C1556124: Unknown - Living Arrangement**
22. **C1556125: Unknown - Transport Arranged**
23. **C1556126: Unknown - Escort Required**
24. **C1556127: Unknown - Patient Outcome**
25. **C1556128: Unknown - Job Status**
26. **C1556129: Unknown - Patient_s Relationship to Insured**
27. **C1556130: Unknown - CWE statuses**
28. **C1556131: Unknown - Container status**
29. **C1556132: Unknown - Immunization Registry Status**
30. **C1556133: Unknown - Expanded yes/no indicator**
31. **C1556134: Unknown - Event Expected**
32. **C1556135: Unknown - Patient Class**
33. **C1556136: Unknown - Living Dependency**

34. **C1556137: Unknown - Contact Role**
35. C1561529: Unknown
36. **C1609613: unknown - NullFlavor**

2.4 "protocols" (degree 24)

Except for 'Protocols documentation', the remaining cases should be marked as suppressible because they mean something more specific than "protocols". The concepts involved are

1. C0442711: Protocols documentation
2. **C0542547: Protocols: Activities**
3. **C0677556: Protocols: Pre- or Intra- or Post-Procedure**
4. **C0677557: Protocols: Urinary Elimination**
5. **C0677558: Protocols: Tissue Perfusion**
6. **C0677559: Protocols: Tissue Integrity**
7. **C0677560: Protocols: Sensation, Pain and Comfort**
8. **C0677561: Protocols: Self-Concept**
9. **C0677562: Protocols: Self-Care**
10. **C0677563: Protocols: Safety**
11. **C0677564: Protocols: Role Relationship**
12. **C0677565: Protocols: Respiration**
13. **C0677566: Protocols: Physical Regulation**
14. **C0677567: Protocols: Nutrition**
15. **C0677568: Protocols: Metabolism**
16. **C0677569: Protocols: Medications and Blood Products**
17. **C0677570: Protocols: Immunology**
18. **C0677571: Protocols: Health Behavior**
19. **C0677572: Protocols: Fluid and Electrolyte**
20. **C0677573: Protocols: Coping**
21. **C0677574: Protocols: Cognition**
22. **C0677575: Protocols: Circulation**
23. **C0677576: Protocols: Bowel Elimination**
24. **C1507394: Clinical trial protocol document**

2.5 "assessment" (degree 20)

Except for 'Assessment procedure' and 'Evaluation', the remaining cases should be suppressed because they are specific kinds of "assessment". The concepts involved in this ambiguity are

1. **C0028708: Nutrition Assessment**
2. **C0031809: Physical Examination**
3. C0220825: Evaluation
4. **C0278371: Physical assessment**
5. **C0542573: Assessment: Bowel Elimination**
6. **C0549068: Assessment: Circulation**
7. **C0549070: Assessment: Coping**
8. **C0549071: Assessment: Fluid and Electrolytes**
9. **C0549072: Assessment: Health Behavior**
10. **C0549073: Assessment: Medications and Blood Products**
11. **C0549074: Assessment: Metabolism**

12. **C0549075: Assessment: Respiration**
13. **C0549076: Assessment: Safety**
14. **C0549077: Assessment: Self-Care**
15. **C0549078: Assessment: Sensation, Pain and Comfort**
16. **C0549079: Assessment: Urinary Elimination**
17. **C0549080: Assessment: Pre- or Intra- or Post-Procedure**
18. **C0679207: Knowledge acquisition using a method of assessment**
19. **C0870300: Assessment: Cognition**
20. C1261322: Assessment procedure

2.6 "ec 2.7.1.112" (degree 19)

All Enzyme Commission (EC) numbers (strings beginning "ec <integer>.") are suppressed by MetaMap because they represent classes of enzymes and are consequently highly ambiguous.

1. *C0033681: Protein Tyrosine Kinase*
2. *C0065344: Lymphocyte Specific Protein Tyrosine Kinase p56(lck)*
3. *C0072470: Proto-Oncogene Protein c-kit*
4. *C0109317: EphB2 Receptor*
5. *C0117718: fibroblast growth factor receptor 3*
6. *C0138965: protein-tyrosine kinase c-src*
7. *C0169658: Janus kinase 1*
8. *C0169661: Janus kinase 2*
9. *C0290067: Platelet-Derived Growth Factor alpha Receptor*
10. *C0290068: Platelet-Derived Growth Factor beta Receptor*
11. *C0907648: Ephrin Receptor EphB1*
12. *C0915156: Ephrin Receptor EphA8*
13. *C1259418: MERTK protein, human*
14. *C1333408: EPHA4 protein, human*
15. *C1333409: EPHB3 protein, human*
16. *C1333410: EPHA2 protein, human*
17. *C1334392: LTK protein, human*
18. *C1370509: EPHA1 protein, human*
19. *C1504624: KDR protein, human*

2.7 "patient education plans" (degree 18) <no change from last year>

All eighteen cases should be suppressed because they are specific kinds of "patient education plans". Their concepts are

1. **C0549081|Patient Education Plans: Activities**
2. **C0549082|Patient Education Plans: Bowel Elimination**
3. **C0549083|Patient Education Plans: Circulation**
4. **C0549084|Patient Education Plans: Coping**
5. **C0549085|Patient Education Plans: Health Behavior**
6. **C0549086|Patient Education Plans: Immunology**
7. **C0549087|Patient Education Plans: Medications and Blood Products**
8. **C0549088|Patient Education Plans: Metabolism**
9. **C0549089|Patient Education Plans: Nutrition**
10. **C0549090|Patient Education Plans: Physical Regulation**

11. **C0549091|Patient Education Plans: Respiration**
12. **C0549092|Patient Education Plans: Role Relationship**
13. **C0549093|Patient Education Plans: Safety**
14. **C0549094|Patient Education Plans: Self-Care**
15. **C0549095|Patient Education Plans: Sensation, Pain and Comfort**
16. **C0549096|Patient Education Plans: Tissue Integrity**
17. **C0549097|Patient Education Plans: Urinary Elimination**
18. **C0549098|Patient Education Plans: Pre- or Intra- or Post-Procedure**

2.8 "emergency" (degree 15)

Except for 'Emergency Situation' and 'Bale out', the remaining cases should be suppressed because they are specific kinds of "emergency". The concepts involved in this ambiguity are

1. C0013956|Emergency Situation
2. C0175673|Bale out
3. **C1546399|Encounter Admission Source - emergency**
4. **C1546844|Visit Priority Code - Emergency**
5. **C1547144|Specialty Type - Emergency**
6. **C1552231|Clinical Nurse Specialist - Emergency**
7. **C1553500|Act Code - emergency**
8. **C1555975|Registered Nurse - Emergency**
9. **C1561583|Patient Class - Emergency**
10. **C1561584|Certification patient type - Emergency**
11. **C1561585|Level of Care - Emergency**
12. **C1561586|Consent Bypass Reason - Emergency**
13. **C1561587|Referral category - Emergency**
14. **C1561588|Admission Type - Emergency**
15. **C1561589|Consent Non-Disclosure Reason - Emergency**

2.9 "none" (degree 13)

Except for 'None', the remaining cases should be suppressed because they are specific kinds of "none". The concepts involved in this ambiguity are

1. C0549184|None
2. **C1546509|none - TableRules**
3. **C1547191|none - ResponseLevel**
4. **C1550083|None - EntityCode**
5. **C1550437|None - Sequencing**
6. **C1551387|None - ContainerSeparator**
7. **C1553523|none - SubstanceAdminSubstitution**
8. **C1556146|None - Relationship**
9. **C1556147|None - Eligibility Source**
10. **C1556148|None - Action Taken in Response to the Event**
11. **C1556150|None - ObservationValue**
12. **C1556151|None - Language Proficiency**
13. **C1556152|None - Additive/Preservative**

2.10 "ec 2.7.1.-" (degree 12) <no change from last year>

All Enzyme Commission (EC) numbers (strings beginning "ec <integer>.") are suppressed by MetaMap because they represent classes of enzymes and are consequently highly ambiguous.

1. *C0108836/CDC7 protein, human*
1. *C0108855/CDK2 protein, human*
1. *C0217310/GRK5 protein, human*
1. *C0258733/GRK6 protein, human*
1. *C0259367/PCTAIRE Protein Kinase 1*
1. *C0659150/CHEK1 protein, human*
1. *C0673406/GPRK7 protein, human*
1. *C1333180/Cyclin-Dependent Kinase 10*
1. *C1333735/GPRK2L protein, human*
1. *C1333738/G Protein-Coupled Receptor Kinase Family*
1. *C1337052/PAK6 protein, human*
1. *C1447440/CDK3 protein, human*

2.11 "active" (degree 11)

Except for 'Active' and 'Active brand of pseudoephedrine-triprolidine', the remaining cases should be suppressed because they are specific kinds of "active". Suppress 'Allergy brand of pseudoephedrine-triprolidine' (MetaMap only) because it is a brand name. The concepts involved in this ambiguity are

1. C0205177|Active
2. *C0718247|Active brand of pseudoephedrine-triprolidine*
3. **C1547419|ActStatus - active**
4. **C1553875|Concept Status - Active**
5. **C1561507|EditStatus - Active**
6. **C1561508|Managed Participation Status - active**
7. **C1561509|Role Status - active**
8. **C1561510|Entity Status - active**
9. **C1561511|Document Storage - active**
10. **C1561512|Document Storage Status - Active**
11. **C1561513|Immunization Registry Status - Active**

2.12 "cat" (degree 10)

Suppress 'Family Felidae' and 'Subfamily Felinae' (MetaMap only) because they are too general and are not normal usage in biomedical literature. Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. C0007450|Felis catus
2. *C0008169/Chloramphenicol O-Acetyltransferase*
3. *C0040405/X-Ray Computed Tomography*
4. *C0280589/cytarabine/thioguanine*
5. *C0325089/Family Felidae*
6. C0325090|Felis silvestris
7. C0524517|Genus Felis
8. *C0683312/Categories*

9. *C1270185|Subfamily Felinae*
10. *C1366498|Chloramphenicol Acetyl Transferase Gene*

2.13 "patient" (degree 10)

Except for 'Patients', the remaining cases should be suppressed because they are specific kinds of "patient". The concepts involved in this ambiguity are

1. C0030705|Patients
2. **C1550655|Specimen Type - Patient**
3. **C1578478|Role Class - patient**
4. **C1578479|Role Code - Patient receipt**
5. **C1578480|Role Code - Patient specimen**
6. **C1578481|Mail Claim Party - Patient**
7. **C1578483|Report source - Patient**
8. **C1578484|Relationship modifier - Patient**
9. **C1578485|Specimen Source Codes - Patient**
10. **C1578486|Disabled Person Code - Patient**

2.14 "radiology" (degree 10)

Except for 'Radiology Speciality', 'Diagnostic radiologic examination' and 'Radiology studies', the remaining cases should be suppressed because they are specific kinds of "radiology". The concepts involved in this ambiguity are

1. C0034599|Radiology Specialty
2. C0043299|Diagnostic radiologic examination
3. C0807679|Radiology studies
4. **C1405978|Encounter due to radiological examination**
5. **C1548000|Radiology Section ID**
6. **C1548429|radiology referral type**
7. **C1552284|Radiology Podiatrist**
8. **C1555923|Radiology Chriopractor** (sic)
9. **C1608525|Radiology - NUCCProvider Codes**
10. **C1610162|Radiology - Clinic/Center - NUCCProviderCodes**

2.15 "sports medicine" (degree 10)

Except for 'sports medicine specialty', the remaining cases should be suppressed because they are specific kinds of "sports medicine specialty". The concepts involved in this ambiguity are

1. C0038040|sports medicine specialty
2. **C1552285|Podiatrist - Sports Medicine**
3. **C1555741|Emergency Medicine - Sports Medicine**
4. **C1555748|Family Practice - Sports Medicine**
5. **C1555771|Internal Medicine - Sports Medicine**
6. **C1555800|Orthopedic Surgery - Sports Medicine**
7. **C1555844|Pediatrics - Sports Medicine**
8. **C1555849|Physical Medicine & Rehabilitation - Sports Medicine**
9. **C1555858|Preventive Medicine - Sports Medicine**

10. **C1555872|Psychiatry & Neurology - Sports Medicine**

2.16 "cam" (degree 9)

Suppress ambiguous form(s) (MetaMap only) because they are either abbreviatory or, in the case of 'CAM brand of Ephedrine Hydrochloride', a brand name. The concepts involved are

1. *C0007578|Cell Adhesion Molecules*
2. *C0054551|cyclophosphamide/doxorubicin/methotrexate protocol*
3. *C0178551|chorioallantoic membrane*
4. *C0678112|CAM brand of Ephedrine Hydrochloride*
5. C0713465|Cam, topical lotion
6. *C1148475|Complementary and alternative medicine*
7. *C1366433|CCM1 gene*
8. *C1366910|Calmodulin 1*
9. *C1366911|Cerebral Cavernous Malformations 1*

2.17 "ec 1.14.14.1" (degree 9) <no change from last year>

All Enzyme Commission (EC) numbers (strings beginning "ec <integer>.") are suppressed by MetaMap because they represent classes of enzymes and are consequently highly ambiguous.

1. *C0003805|Aromatase*
2. *C0010763|Cytochrome P-450 CYP2E1*
3. *C0057223|Cytochrome P-450 CYP2D6*
4. *C0059735|Cytochrome P-450 CYP1A1*
5. *C0207509|Cytochrome P-450 CYP1A2*
6. *C0608437|CYP2C9 protein, human*
7. *C0960580|CYP2C19 protein, human*
8. *C1142644|Cytochrome P450 3A4*
9. *C1321893|cytochrome P-450 CYP2C18*

2.18 "ec 3.4.22.-" (degree 9) <no change from last year>

All Enzyme Commission (EC) numbers (strings beginning "ec <integer>.") are suppressed by MetaMap because they represent classes of enzymes and are consequently highly ambiguous.

1. *C0291573|caspase-3*
2. *C0537969|caspase-7*
3. *C0538091|caspase-6*
4. *C0758959|caspase-10*
5. *C0769345|caspase-5*
6. *C0910167|caspase-9*
7. *C0912403|cathepsin Z*
8. *C1437927|CASP4 protein, human*
9. *C1453171|CASP2 protein, human*

2.19 "microsomal monooxygenase" (degree 9) <no change from last year>

All nine cases should be suppressed because they are specific kinds of "microsomal monooxygenases". Their concepts are

1. **C0003805|Aromatase**
2. **C0003927|Aryl Hydrocarbon Hydroxylases**
3. **C0010763|Cytochrome P-450 CYP2E1**
4. **C0057223|Cytochrome P-450 CYP2D6**
5. **C0059735|Cytochrome P-450 CYP1A1**
6. **C0207509|Cytochrome P-450 CYP1A2**
7. **C0608437|CYP2C9 protein, human**
8. **C0960580|CYP2C19 protein, human**
9. **C1321893|cytochrome P-450 CYP2C18**

2.20 "no" (degree 9)

Except for 'no' and No - yes/no indicator', the remaining cases should be suppressed because they are specific kinds of "no". The concepts involved in this ambiguity are

1. C1298908|no
2. **C1546943|No - Event Seriousness**
3. **C1546946|No - Event Expected**
4. **C1546967|No - Identity May Be Divulged**
5. **C1548170|No - Release Information**
6. **C1549056|No - Expanded yes/no indicator**
7. **C1549062|No - Notify Clergy Code**
8. **C1549442|No - Assignment of Benefits**
9. C1549444|No - yes/no indicator

2.21 "not applicable" (degree 9)

Except for 'not applicable', the remaining cases should be suppressed because they are specific kinds of "not applicable". The concepts involved in this ambiguity are

1. C1272460|not applicable
2. **C1546968|No Information - not applicable**
3. **C1547280|Production Class Code - Not Applicable**
4. **C1549103|Administrative Sex - Not applicable**
5. **C1609491|Patient Class - Not Applicable**
6. **C1610044|Derived specimen - Not Applicable**
7. **C1610595|Identity May Be Divulged - Not applicable**
8. **C1611147|CWE statuses - Not applicable**
9. **C1619691|Expanded yes/no indicator - not applicable**

2.22 "p" (degree 9)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0033095|Pressure- physical agent*
2. *C0033452|Properdin*

3. *C0202178|Phosphorus measurement*
4. *C0227095|Deciduous mandibular right central incisor tooth*
5. *C0391850|Physiologic pulse*
6. *C0439115|upper case pea*
7. *C0439140|lower case pea*
8. *C1301732|Planned*
9. *C1521733|Pathologic*

2.23 "pediatrics" (degree 9)

Except for 'Pediatric discipline', the remaining cases should be suppressed because they are specific kinds of "pediatrics". The concepts involved in this ambiguity are

1. C0030755|Pediatric discipline
2. **C1552244|Clinical Nurse Specialist - Pediatrics**
3. **C1552268|Nurse Practitioner - Pediatrics**
4. **C1552293|Occupational Therapist - Pediatrics**
5. **C1552305|Physical Therapist - Pediatrics**
6. **C1555461|Act Code - Pediatrics**
7. **C1555826|Physicians - Pediatrics**
8. **C1555953|Optometrist - Pediatrics**
9. **C1556004|Registered Nurse - Pediatrics**

2.24 "yes" (degree 9)

Except for 'Yes - Yes/no indicator', the remaining cases should be suppressed because they are specific kinds of "Yes". The concepts involved in this ambiguity are

1. **C1298907|Yes - Presence findings**
2. **C1546945|Yes - Event Seriousness**
3. **C1546947|Yes - Event Expected**
4. **C1546969|Yes - Identity May Be Divulged**
5. **C1548171|Yes - Release Information**
6. **C1549060|Yes - Expanded yes/no indicator**
7. **C1549065|Yes - Notify Clergy Code**
8. **C1549443|Yes - Assignment of Benefits**
9. C1549445|Yes - Yes/no indicator

2.25 "allergy" (degree 8)

Suppress 'Allergy brand of chlorpheniramine' (MetaMap only) because it is a brand name. 'Allergy to Causative Agent Code', 'Allergy - Charge Type Reaction', and 'Allergy - Physicians - NUCCProviderCodes' should be suppressed because they are specific kinds of "allergy". The concepts involved are

1. C0002111|Allergy Specialty
2. C0020517|Hypersensitivity
3. C0020523|Immediate hypersensitivity
4. *C0718383|Allergy brand of chlorpheniramine*
5. C1314973|Response to antigens

6. **C1547315|Allergy to Causative Agent Code**
7. **C1547542|Allergy - Charge Type Reason**
8. **C1611280|Allergy - Physicians - NUCCProviderCodes**

2.26 "cde genotype" (degree 8) <no change from last year>

The eight strings "CDE genotype", "CDe genotype", …, "cde genotype" differ only in the case of "c", "d" and "e". They seem to be legitimate concepts from the Read Codes. In a case sensitive environment, there is no ambiguity. But when case is ignored, as in most NLP applications, the ambiguity exists. Nevertheless, the ambiguous forms are correctly not suppressed. The concepts involved are

1. C0427651|CDE genotype
2. C0427652|CDe genotype
3. C0427653|CdE genotype
4. C0427654|Cde genotype
5. C0427655|cDE genotype
6. C0427656|cDe genotype
7. C0427657|cdE genotype
8. C0427658|cde genotype

2.27 "cde haplotype" (degree 8) <no change from last year>

This case of ambiguity is entirely analogous to the "cde genotype" case. The concepts involved are

1. C1319020|cde haplotype finding
2. C1319021|Cde haplotype finding
3. C1319022|cdE haplotype finding
4. C1319023|CDe haplotype finding
5. C1319024|cDE haplotype finding
6. C1319025|cDe haplotype finding
7. C1319026|CdE haplotype finding
8. C1319027|CDE haplotype finding

2.28 "cde haplotype (finding)" (degree 8) <no change from last year>

This case of ambiguity is also entirely analogous to the "cde haplotype" case with the same concepts.

1. C1319020|cde haplotype finding
2. C1319021|Cde haplotype finding
3. C1319022|cdE haplotype finding
4. C1319023|CDe haplotype finding
5. C1319024|cDE haplotype finding
6. C1319025|cDe haplotype finding
7. C1319026|CdE haplotype finding
8. C1319027|CDE haplotype finding

2.29 "cde haplotype finding" (degree 8) <no change from last year>

This case of ambiguity is entirely analogous to the "cde haplotype" case and involves the same concepts.

1. C1319020|cde haplotype finding
2. C1319021|Cde haplotype finding
3. C1319022|cdE haplotype finding
4. C1319023|CDe haplotype finding
5. C1319024|cDE haplotype finding
6. C1319025|cDe haplotype finding
7. C1319026|CdE haplotype finding
8. C1319027|CDE haplotype finding

2.30 "driver injured in collision with other and unspecified motor vehicles in nontraffic accident" (degree 8) <no change from last year>

These strings occur in the Read Codes and ICD-10 with some variation in the wording. Except for the last concept below, they should be suppressed because they involve a more specific kind of driver. The concepts involved are

1. **C0476822|Driver pedal cycle injured in collision with other and unspecified motor vehicles in nontraffic accident**
2. **C0476905|Motorcycle driver injured in collision with other and unspecified motor vehicles in nontraffic accident**
3. **C0476973|Car driver injured in collision with other and unspecified motor vehicles in nontraffic accident**
4. **C0477136|Bus driver injured in collision with other and unspecified motor vehicles in nontraffic accident**
5. **C0496239|Driver of three-wheeled motor vehicle injured in collision with other and unspecified motor vehicles in nontraffic accident**
6. **C0496408|Pick-up truck or van driver injured in collision with other and unspecified motor vehicles in nontraffic accident**
7. **C0496481|Heavy transport vehicle driver injured in collision with other and unspecified motor vehicles in nontraffic accident**
8. C0596026|Driver injured in collision with other and unspecified motor vehicles in nontraffic accident

2.31 "driver injured in collision with other and unspecified motor vehicles in traffic accident" (degree 8) <no change from last year>

This case is analogous to the previous one (section 2.30); all but the last should be suppressed. The concepts involved are

1. **C0476826|Pedal cycle driver injured in collision with other and unspecified motor vehicles in traffic accident**
2. **C0476909|Motor cycle driver injured in collision with other and unspecified motor vehicles in traffic accident**
3. **C0476977|Car driver injured in collision with other and unspecified motor vehicles in traffic accident**

4. **C0477140|Bus driver injured in collision with other and unspecified motor vehicles in traffic accident**
5. **C0496241|Three-wheeled motor vehicle driver injured in collision with other and unspecified motor vehicles in traffic accident**
6. **C0496410|Pick-up truck or van driver injured in collision with other and unspecified motor vehicles in traffic accident**
7. **C0496483|Heavy transport driver injured in collision with other and unspecified motor vehicles in traffic accident**
8. C0596027|Driver injured in collision with other and unspecified motor vehicles in traffic accident

2.32 "ec 2.7.1.37" (degree 8)

All Enzyme Commission (EC) numbers (strings beginning "ec <integer>.") are suppressed by MetaMap because they represent classes of enzymes and are consequently highly ambiguous.

1. *C0033640|PROTEIN KINASE*
2. *C0072402|Protein-Serine-Threonine Kinases*
3. *C0294209|LIM Domain Kinase 1*
4. *C0380146|activin receptor-like kinase 1*
5. *C0541150|3-Phosphoinositide Dependent Protein Kinase-1*
6. *C1314894|Col4A3 protein, human*
7. *C1332856|Casein Kinase 2, Alpha 1 Polypeptide*
8. *C1447968|ACVR1 protein, human*

2.33 "flavoprotein-linked monooxygenase" (degree 8) <no change from last year>

This case is analogous to "microsomal monooxygenase" (section 2.19). All eight cases should be suppressed because they are specific kinds of "flavoprotein-linked monooxygenases". Their concepts are

1. **C0003805|Aromatase**
2. **C0010763|Cytochrome P-450 CYP2E1**
3. **C0057223|Cytochrome P-450 CYP2D6**
4. **C0059735|Cytochrome P-450 CYP1A1**
5. **C0207509|Cytochrome P-450 CYP1A2**
6. **C0608437|CYP2C9 protein, human**
7. **C0960580|CYP2C19 protein, human**
8. **C1321893|cytochrome P-450 CYP2C18**

2.34 "mac" (degree 8)

Suppress ambiguous form(s) (MetaMap only) because they are either abbreviatory or, in the case of 'Raincoat', slang usage. The concepts involved are

1. *C0009545|Complement Membrane Attack Complex*
2. *C0026916|Mycobacterium avium-intracellulare Infection*
3. *C0065465|cyclophosphamide/dactinomycin/methotrexate protocol*
4. *C0083360|chlorambucil/dactinomycin/methotrexate protocol*
5. *C0279190|cyclophosphamide/doxorubicin/mitomycin protocol*

6. *C0451273|MacAndrew Alcoholism Scale*
7. *C0453947|Raincoat*
8. *C1167383|membrane attack complex location*

2.35 "normal" (degree 8)

Except for 'Normal', the remaining cases should be suppressed because they are specific kinds of "normal". The concepts involved in this ambiguity are

1. C0205307|Normal
2. **C1550457|Normal Observation Interpretation**
3. **C1550469|normal Confidentiality**
4. **C1551394|normal Device Alert Level**
5. **C1553386|normal Act Status**
6. **C1553399|normal Managed Participation Status**
7. **C1553402|normal Role Status**
8. **C1553406|normal Entity Status**

2.36 "passenger injured in collision with other and unspecified motor vehicles in nontraffic accident" (degree 8)

This case is analogous to the ones in section 2.30 and 2.31; all but the last should be suppressed. The concepts involved are

1. **C0476823|Pedal cycle passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
2. **C0476906|Motor cycle passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
3. **C0496240|Three-wheeled motor vehicle passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
4. **C0496324|Passenger of car injured in collision with other and unspecified motor vehicles in nontraffic accident**
5. **C0496409|Pick-up truck or van passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
6. **C0496482|Heavy transport vehicle passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
7. **C0496486|Bus passenger injured in collision with other and unspecified motor vehicles in nontraffic accident**
8. C1540954|Passenger injured in collision with other and unspecified motor vehicles in nontraffic accident

2.37 "rehabilitation" (degree 8)

Except for 'Rehabilitation therapy', 'Rehabilitation aspects' and 'Rehabilitation - specialty', the remaining cases should be suppressed because they are specific kinds of "rehabilitation". The concepts involved in this ambiguity are

1. **C0007237|Encounter due to care involving use of rehabilitation procedures**
2. C0034991|Rehabilitation therapy
3. C0034992|Rehabilitation aspects

4. C1306847|Rehabilitation - specialty
5. **C1552254|Clinical Nurse Specialist - Rehabilitation**
6. **C1552452|Clinic/Center - Rehabilitation**
7. **C1555915|Psychologist - Rehabilitation**
8. **C1556010|Registered Nurse - Rehabilitation**

2.38 "routine" (degree 8)

Except for 'Routine', the remaining cases should be suppressed because they are specific kinds of "routine". The concepts involved in this ambiguity are

1. C0205547|Routine
2. **C1546402|Admission Type - Routine**
3. **C1546438|Priority - Routine**
4. **C1547137|Level of Care - Routine**
5. **C1547582|Extended Priority Codes - Routine**
6. **C1548424|Referral priority - Routine**
7. **C1549563|Processing priority - Routine**
8. **C1561591|Report priority - Routine**

2.39 "school" (degree 8)

Except for 'School (environment)', the remaining cases should be suppressed because they are specific kinds of "school". The concepts involved in this ambiguity are

1. C0036375|School (environment)
2. **C1552255|School - Clinical Nurse Specialist - NUCCProviderCodes**
3. **C1552273|School - Nurse Practitioner - NUCCProviderCodes**
4. **C1555917|School - Social Worker - NUCCProviderCodes**
5. **C1556012|School - Registered Nurse - NUCCProviderCodes**
6. **C1608526|school - Role Code**
7. **C1617581|School - Psychologist - NUCCProviderCodes**
8. **C1618606|School - Counselor - NUCCProviderCodes**

2.40 "t" (degree 8)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0039476|Temperature*
2. *C0040715|Chromosomal translocation*
3. *C0227099|Mandibular right second primary molar*
4. *C0439119|Upper case tea*
5. *C0439143|Lower case tea*
6. *C0475455|T - Tumor stage*
7. *C1522168|Topical Route of Drug Administration*
8. *C1552647|ProbabilityDistributionType - T*

2.41 "a" (degree 7)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0004398|Autopsy*
2. *C0205146|Area*
3. *C0227089|Deciduous maxillary right second molar tooth*
4. *C1442985|a - tumor staging*
5. *C1442986|A - tumor staging*
6. *C1522424|A Mouse*
7. *C1535245|A*

2.42 "ar" (degree 7)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0051755|Amphiregulin*
2. *C0332284|Arising in*
3. *C0559546|Adverse reactions*
4. *C1332109|AREG gene*
5. *C1367578|AR gene*
6. *C1447749|AR protein, human*
7. *C1514768|Recombinant Amphiregulin*

2.43 "cd32" (degree 7)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0054951|CD32 Antigens*
2. *C0215037|FCGR2C protein, human*
3. *C1414553|FCGR2A gene*
4. *C1414554|FCGR2B gene*
5. *C1448238|FCGR2A protein, human*
6. *C1452607|FCGR2B protein, human*
7. *C1521989|FCGR2C gene*

2.44 "coagulation" (degree 7)

Except for 'Blood coagulation', 'Coagulation procedure' and 'Coagulation process', the remaining cases should be suppressed because they are specific kinds of "rehabilitation". The concepts involved in this ambiguity are

1. C0005778|Blood coagulation
2. **C0005790|Blood coagulation tests**
3. **C0427579|Blood coagulation pathway observation**
4. C0441509|Coagulation procedure
5. C1328723|Coagulation process
6. **C1556326|Adverse Event Associated with Coagulation**
7. **C1561952|Observation Method - Coagulation**

2.45 "critical care medicine" (degree 7)

Except for 'Critical care medicine specialty', the remaining cases should be suppressed because they are specific kinds of "critical care medicine". The concepts involved in this ambiguity are

1. C1319411|Critical care medicine specialty
2. **C1552230|Clinical Nurse Specialist - Critical Care Medicine**
3. **C1552261|Nurse Practitioner - Critical Care Medicine**
4. **C1555727|Physicians - Anesthesiology - Critical Care Medicine**
5. **C1555757|Physicians - Internal Medicine - Critical Care Medicine**
6. **C1555788|Physicians - Obstetrics & Gynecology - Critical Care Medicine**
7. **C1555972|Registered Nurse - Critical Care Medicine**

2.46 "ctf" (degree 7)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0069116|NFI Transcription Factors*
2. *C1417705|NFIC gene*
3. *C1431185|NFIB protein, human*
4. *C1455428|NFIA protein, human*
5. *C1513835|NF-I genes*
6. *C1522145|NFIX protein, human*
7. *C1522146|NFIC protein, human*

2.47 "ice" (degree 7)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory or, in the case of 'Methamphetamine', slang usage. The concepts involved are

1. C0020746|Ice
2. *C0025611|Methamphetamine*
3. *C0249492|cytarabine/etoposide/idarubicin*
4. *C0280697|carboplatin/etoposide/ifosfamide*
5. *C0534519|Caspase-1*
6. *C1366479|CASP1 gene*
7. *C1521869|Ifosfamide/Cyclophosphamide/Etoposide*

2.48 "isolation" (degree 7)

Except for 'Isolation procedure' and 'isolation aspects', the remaining cases should be suppressed because they are specific kinds of "isolation". The concepts involved in this ambiguity are

1. C0204727|Isolation procedure
2. C0220862|isolation aspects
3. **C0260397|Need for isolation**
4. **C1547129|Act Code - Isolation**
5. **C1548355|Specialty Type - Isolation**
6. **C1608299|Privacy Level - Isolation**
7. **C1608300|Level of Care - Isolation**

2.49 "line" (degree 7)

'line source specimen code' and 'Vein, Artery or Body Cavity Line Code' should be suppressed because they are specific kinds of "line". Suppress 'Long Interspersed Elements' (MetaMap only) because it is abbreviatory. The concepts involved in this ambiguity are

1. C0205132|Linear
2. C0700221|Intravascular line
3. *C1517938|Long Interspersed Elements*
4. **C1546701|line source specimen code**
5. C1550648|Line Specimen
6. C1552960|line unit of measure
7. **C1556143|Vein, Artery or Body Cavity Line Code**

2.50 "m" (degree 7)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0024554|Male gender*
2. *C0026473|monocyte*
3. *C0227102|Mandibular left primary canine tooth*
4. *C0439113|Upper case emm*
5. *C0456533|M - Metastasis stages*
6. *C0475209|meter*
7. *C1535257|M NOS*

2.51 "mif" (degree 7)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0024429|Migration Inhibitory Factor*
2. *C0054504|Calgranulin A*
3. *C0054505|Calgranulin B*
4. *C0312359|phenylpyruvate tautomerase*
5. *C1334507|MIF gene*
6. *C1335798|S100A9 gene*
7. *C1366582|S100A8 gene*

2.52 "nutrition" (degree 7)

'Nutrition aspect of cancer' and 'Chiropractor - Nutrition' should be suppressed because they are specific kinds of "nutrition". The concepts involved are

1. C0028707|Science of nutrition
2. C0518896|Nutrition outcomes
3. C0600072|Feeding and dietary regimes
4. C1442959|Nutrition function
5. **C1443053|Nutrition aspect of cancer**
6. C1521729|Nutritional Study
7. **C1555920|Chiropractor - Nutrition**

2.53 "passenger injured in collision with other and unspecified motor vehicles in traffic accident" (degree 7) <no change from last year>

This case is analogous to 2.36; all ambiguous strings except for the third sense should be suppressed. The concepts involved are

1. **C0476827|Pedal cycle passenger injured in collision with other and unspecified motor vehicles in traffic accident**
2. **C0476910|Motor cycle passenger injured in collision with other and unspecified motor vehicles in traffic accident**
3. C0496242|Passenger injured in collision with other and unspecified motor vehicles in traffic accident
4. **C0496326|Passenger of car injured in collision with other and unspecified motor vehicles in traffic accident**
5. **C0496411|Passenger of pick-up truck or van vehicle injured in collision with other and unspecified motor vehicles in traffic accident**
6. **C0496484|Passenger of heavy transport vehicle vehicle injured in collision with other and unspecified motor vehicles in traffic accident**
7. **C0496488|Passenger of bus injured in collision with other and unspecified motor vehicles in traffic accident**

2.54 "peroxisome vesicle" (degree 7)

Except for 'peroxisome', the remaining cases should be suppressed because they are specific kinds of "peroxisome vesicle". The concepts involved in this ambiguity are

1. C0752063|peroxisome
2. **C1166933|P1 peroxisome**
3. **C1166934|P2 peroxisome**
4. **C1166935|P3 peroxisome**
5. **C1166936|P4 peroxisome**
6. **C1166937|P5 peroxisome**
7. **C1166938|P6 peroxisome**

2.55 "stat" (degree 7)

'Extended Priority Codes - Stat', 'Referral priority - STAT' and 'Report priority - Stat' should be suppressed because they are specific kinds of "stat". Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved in this ambiguity are

1. C0205548|Stat (do immediately)
2. *C0449438|Status*
3. *C1335960|STAT family gene*
4. *C1366753|STAT protein*
5. **C1547583|Extended Priority Codes - Stat**
6. **C1548425|Referral priority - STAT**
7. **C1561443|Report priority - Stat**

2.56 "urgent" (degree 7)

Except for 'Urgent', the remaining cases should be suppressed because they are specific kinds of "urgent". The concepts involved in this ambiguity are

1. C0439609|Urgent
2. **C1546403|Admission Type - Urgent**
3. **C1546845|Visit Priority Code - Urgent**
4. **C1547230|Triage Code - Urgent**
5. **C1561553|Act Priority - urgent**
6. **C1561555|Act Code - urgent**
7. **C1561556|Certification patient type - Urgent**

2.57 "adh" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0001942|Alcohol dehydrogenase*
2. *C0003779|Argipressin*
3. *C0042413|Vasopressins*
4. *C0201849|Antidiuretic hormone measurement*
5. *C1332347|Atypical Ductal Breast Hyperplasia*
6. *C1366535|AVP gene*

2.58 "all" (degree 6)

'all - LocalMarkupIgnore' and 'all - TableRules' should be suppressed because they are specific kinds of "all". Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0023449|Leukemia, Lymphocytic, Acute*
2. *C0105421|BCR protein, human*
3. C0444868|All
4. *C0812385|BCR gene*
5. **C1552816|all - LocalMarkupIgnore**
6. **C1552837|all - TableRules**

2.59 "apc" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0003315|Antigen-Presenting Cells*
2. *C0033036|Atrial Premature Complexes*
3. *C0162832|APC gene*
4. *C0753723|Adenomatous Polyposis Coli Protein*
5. *C1176286|anaphase-promoting complex location*
6. *C1510852|Alkylphosphocholine Compound*

2.60 "bar" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0001643|beta-2 Adrenergic Receptors*
2. C0441233|External fixator bar
3. C0687760|Taverns
4. C0993613|Bar form
5. *C1367657|ADRB2 gene*
6. C1551065|bar unit of measure

2.61 "bone" (degree 6)

'Specimen Source Codes - Bone' and 'Specimen Type - Bone' should be suppressed because they are specific kinds of "bone". The concepts involved are

1. C0262950|Skeletal bone
2. C0391978|Bone Tissue
3. C1266908|Entire bone (organ)
4. C1266909|Entire bony skeleton
5. **C1546560|Specimen Source Codes - Bone**
6. **C1550616|Specimen Type - Bone**

2.62 "ca" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0006675|Calcium*
2. *C0006826|Malignant Neoplasms*
3. *C0007097|Carcinoma*
4. *C0019564|Hippocampus (Brain)*
5. *C0050385|cyclophosphamide/doxorubicin protocol*
6. *C0201925|Calcium measurement*

2.63 "cd" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0006632|Cadmium*
2. *C0056447|CP protocol*
3. *C0079141|Compact discs*
4. *C0332140|Diagnosis, clinical*
5. *C0700300|candela*
6. *C0805701|Code*

2.64 "cdr" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0011485/Deoxycytidine*
2. *C0021024/Complementarity Determining Regions*
3. *C0055351/Chlordecone reductase*
4. *C1332741/CDR1 gene*
5. *C1366637/CBFA2T1 gene*
6. *C1437277/CDR1 protein, human*

2.65 "cl" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0010575/Cycloleucine*
2. *C0201952/Chloride measurement*
3. *C0338272/cyclophosphamide/losoxantrone*
4. *C0475212/centiliter*
5. *C0596019/Chloride Ion*
6. *C1521802/Closure*

2.66 "cold" (degree 6) <no change from last year>

The concepts associated with the six senses of "cold" are

1. C0009264|cold temperature
2. C0009443|Common Cold
3. **C0010412|Cold Therapy** [consider suppressing]
4. *C0024117/Chronic Obstructive Airway Disease*
5. C0234192|Cold Sensation
6. *C0719425/Cold brand of chlorpheniramine-phenylpropanolamine*

Senses 3, 4 and 6 are currently suppressed by MetaMap for mainly practical reasons. The only sense that might be marked as suppressible in the Metathesaurus is sense 4, 'Cold Therapy', because it is extremely rare to find the word "cold" by itself meaning 'Cold Therapy'.

2.67 "critical" (degree 6)

Except for 'Critical', the remaining cases should be suppressed because they are specific kinds of "critical". The concepts involved in this ambiguity are

1. C1511545|Critical
2. **C1547298|Admission Level of Care Code - Critical**
3. **C1547308|Patient Condition Code - Critical**
4. **C1551396|DeviceAlertLevel - Critical**
5. **C1552746|EquipmentAlertLevel - Critical**
6. **C1579038|Alert level - Critical**

2.68 "d" (degree 6)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0227086|Maxillary right lateral primary incisor*
2. *C0332173|Daily*
3. *C0439107|Upper case dee*
4. *C0439129|Lower case dee*
5. *C0439228|day*
6. *C1328653|D (tumor staging)*

2.69 "dental" (degree 6)

Except for 'Dental', the remaining cases should be suppressed because they are specific kinds of "dental". The concepts involved in this ambiguity are

1. **C0011344|Dental Clinics**
2. C0226984|Dental
3. **C1522187|Dental Route of Drug Administration**
4. **C1547182|School type - Dental**
5. **C1550246|Body Parts - Dental**
6. **C1561520|Act Code - Dental**

2.70 "depression" (degree 6) <no change from last year>

Suppress 'Cancer patients and suicide and depression' because it is more specific than "depression". The concepts involved are

1. C0011570|Mental Depression
2. C0011581|Depressive disorder
3. C0344315|Depressed mood
4. C0349217|Depressive episode, unspecified
5. C0460137|Depression motion
6. **C0812393|Cancer patients and suicide and depression**

2.71 "education" (degree 6)

Suppress 'Educational Status', 'Details of education' and 'Specialty Type - Education' because they are more specific than "education". The concepts involved are

1. C0013621|Knowledge acquisition
2. C0013622|Educational aspects
3. **C0013658|Educational Status**
4. C0039401|Educational process of instructing
5. **C0424927|Details of education**
6. **C1547143|Specialty Type - Education**

2.72 "emergency care" (degree 6)

Suppress 'Respiratory Therapist, Certified - Emergency Care', 'Respiratory Therapist, Registered - Emergency Care' and 'Clinic / Center - Emergency Care' because they are more specific than "emergency care". The concepts involved in this ambiguity are

1. C0013961|Emergency medical service
2. C0013969|Emergency treatment
3. C1527398|Emergency Care
4. **C1552317|Respiratory Therapist, Certified - Emergency Care**
5. **C1552331|Respiratory Therapist, Registered - Emergency Care**
6. **C1552423|Clinic / Center - Emergency Care**

2.73 "employee" (degree 6)

Except for 'employee' (occurs twice), the remaining cases should be suppressed because they are specific kinds of "employee". The concepts involved in this ambiguity are

1. C0599987|Employee
2. **C1546499|Relationship - Employee**
3. **C1548657|Employee - Patients Relationship to Insured**
4. **C1549077|Employee - Institution Relationship Type**
5. **C1549590|Employee - Confidentiality code**
6. C1619827|employee

2.74 "ent" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0150934|Ear, nose and throat*
2. *C0260087|Otolaryngologist*
3. *C0439751|Entire*
4. *C1274038|Ear, nose and throat surgery*
5. *C1304535|Entitic*
6. *C1521975|Entered*

2.75 "fat" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0015677|Fatty acid glycerol esters*
2. *C0279453|doxorubicin/fluorouracil/triazinate protocol*
3. C0424612|Obese build
4. *C0812278|FAT gene*
5. *C1366645|CD36 gene*
6. *C1435181|FAT protein, human*

2.76 "general practice" (degree 6)

Except for 'general practice (field)', the remaining cases should be suppressed because they are specific kinds of "general practice". The concepts involved in this ambiguity are

1. C0086343|general practice (field)
2. **C1552281|Podiatrist - General Practice**
3. **C1555749|Allopathic & Osteopathic Physicians - General Practice**
4. **C1555929|Dentist - General Practice**
5. **C1555979|Registered Nurse - General Practice**
6. **C1556040|Pharmacist - General Practice**

2.77 "hematology" (degree 6)

Except for 'Hematology' and 'Hematology procedure', the remaining cases should be suppressed because they are specific kinds of "hematology". The concepts involved in this ambiguity are

1. C0018943|Hematology
2. C0200627|Hematology procedure
3. **C1547985|Diagnostic Service Section ID - Hematology**
4. **C1552392|Specialist / Technologist, Pathology - Hematology**
5. **C1555761|Physicians - Internal Medicine - Hematology**
6. **C1555820|Physicians - Pathology - Hematology**

2.78 "hydroxyaryl-protein kinase" (degree 6)

All strings except for the first one are more specific than "hydroxyaryl-protein kinase" and should be suppressed. The concepts involved are

1. C0033681|Protein Tyrosine Kinase
2. **C0117718|fibroblast growth factor receptor 3**
3. **C0215512|RYK protein, human**
4. **C0250935|MATK protein, human**
5. **C0915156|Ephrin Receptor EphA8**
6. **C0919475|RET Protein**

2.79 "k" (degree 6)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0032821|Potassium*
2. *C0202194|Potassium measurement*
3. *C0206065|Karnofsky Performance Status*
4. *C0227104|Deciduous mandibular left second molar tooth*
5. *C0439112|Upper Case Kay*
6. *C0439137|Lower case kay*

2.80 "l" (degree 6)

All single alphabetic strings are suppressed by MetaMap because they are highly ambiguous.

1. *C0023884|Liver*

2. *C0227103|Mandibular left first primary molar*
3. *C0441994|Lower*
4. *C0475211|liter*
5. *C1444754|Length*
6. *C1522275|Lansky Play-Performance Status*

2.81 "mace" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0280578|cytarabine/etoposide/methotrexate protocol*
2. *C0286421|ACE protocol 2*
3. C0349381|Mace Spice
4. *C0813802|Mace brand of omega-Chloroacetophenone*
5. C0949745|Myristica fragrans
6. C1446581|Mace antigen

2.82 "mvp" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0026267|Mitral Valve Prolapse*
2. *C0084989|cisplatin/etoposide/methotrexate chemotherapy protocol*
3. *C0280017|medroxyprogesterone/mitomycin/vinblastine*
4. *C0280018|procarbazine/semustine/vincristine protocol*
5. *C0280540|cisplatin/mitomycin/vinblastine protocol*
6. *C0280541|cisplatin/mitomycin/vindesine protocol*

2.83 "pi3k" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0044602|1-Phosphatidylinositol 3-Kinase*
2. *C0285761|Phosphoinositide-3-Kinase, Catalytic, Gamma Polypeptide*
3. *C1150481|PI3K activity*
4. *C1335214|PIK3CG gene*
5. *C1368105|Phosphoinositide-3-Kinase, Catalytic, Beta Polypeptide*
6. *C1451005|PIK3CA protein, human*

2.84 "premolar tooth" (degree 6) <no change from last year>

All strings except for the last one are more specific than "premolar tooth" and should be suppressed. The concepts involved are

1. **C0005373|Bicuspid**
2. **C0447259|Entire maxillary right first premolar tooth**
3. **C0447265|Permanent upper left second premolar tooth**
4. **C0447266|Entire maxillary left first premolar tooth**

5. **C0447299|Entire mandibular right second premolar tooth**
6. C1281668|Entire premolar tooth

2.85 "pt" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0030705|Patients*
2. *C0032207|Platinum*
3. *C0175252|Paratenial Nucleus*
4. *C0560012|pint*
5. *C0949766|Physical therapy*
6. *C1442880|Point in time*

2.86 "ptc" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0015491|Factor IX*
2. *C0203085|Percutaneous transhepatic cholangiography*
3. *C0694887|PTCH gene*
4. *C0694890|RET gene*
5. *C1366464|F9 gene*
6. *C1425774|CCDC6 gene*

2.87 "rad" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0034599|Radiology Specialty*
2. *C0249069|RAS-related associated with diabetes protein*
3. *C0541315|everolimus*
4. *C0556643|radiation absorbed dose*
5. *C0807679|Radiology studies*
6. *C0919455|RRAD gene*

2.88 "reg" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0205272|Regular*
2. *C0557514|Registrar*
3. *C1367595|EXTL3 gene*
4. *C1514821|Registration*
5. *C1533694|Registering*
6. *C1578536|Registered*

2.89 "set" (degree 6)

Suppress 'Parameterized Data Type - Set' and 'Data Type - ANY - Set' because they are more specific than "set". Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved in this ambiguity are

1. C0036849|Set (Psychology)
2. *C1137947|SET protein, human*
3. *C1366517|SET gene*
4. C1442518|Set scale
5. **C1552652|Parameterized Data Type - Set**
6. **C1552685|Data Type - ANY - Set**

2.90 "spec" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0037778|Medical Specialities*
2. *C0205369|Specific qualifier value*
3. *C0205555|Special (qualifier)*
4. *C0370003|Specimen*
5. *C1315074|Specimen characteristics*
6. *C1521902|To specify*

2.91 "surgery" (degree 6)

Suppress 'Surgeon', 'Surgical procedure finding' and 'Level of Care - Surgery' because they are more specific than "surgery". The concepts involved in this ambiguity are

1. C0038894|Surgery specialty
2. C0038895|Surgical aspects
3. C0543467|Operative Surgical Procedures
4. **C0582175|Surgeon**
5. **C1457907|Surgical procedure finding**
6. **C1547138|Level of Care - Surgery**

2.92 "tumor necrosis factor" (degree 6)

Except for 'Tumor Necrosis Factors', the remaining cases should be suppressed because they are specific kinds of "tumor necrosis factor". The concepts involved in this ambiguity are

1. C0041368|Tumor Necrosis Factors
2. **C1323363|tumor necrosis factor receptor binding**
3. **C1448177|TNF protein, human**
4. **C1456820|Tumor Necrosis Factor-alpha**
5. **C1522668|Recombinant Tumor Necrosis Factor Family Protein**
6. **C1522669|Recombinant Tumor Necrosis Factor-Alpha**

2.93 "vac" (degree 6)

Suppress ambiguous form(s) (MetaMap only) because they are abbreviatory. The concepts involved are

1. *C0042196|Vaccination*
2. *C0042210|Vaccines*
3. *C0054889|cyclophosphamide/doxorubicin/vincristine protocol*
4. *C0077964|cyclophosphamide/dactinomycin/vincristine protocol*
5. *C0280062|cyclophosphamide/doxorubicin/vinblastine protocol*
6. *C0280063|cyclophosphamide/dactinomycin/vinblastine protocol*

2.94 "vitamin b6 metabolism" (degree 6)

Except for 'vitamin B6 metabolism', the remaining cases should be suppressed because they are specific kinds of "vitamin b6 metabolism". The concepts involved in this ambiguity are

1. **C1158215|pyridoxine metabolism**
2. **C1327183|pyridoxal phosphate metabolism**
3. C1327397|vitamin B6 metabolism
4. **C1327398|pyridoxal metabolism**
5. **C1327399|pyridoxamine metabolism**
6. **C1527414|Vitamin B6 Metabolism Pathway**

2.95 "xenobiotic monooxygenase" (degree 6) <no change from last year>

All strings except for the last one are more specific than "xenobiotic monooxygenase" and should be suppressed. The concepts involved are

1. **C0010763|Cytochrome P-450 CYP2E1**
2. **C0057223|Cytochrome P-450 CYP2D6**
3. **C0207509|Cytochrome P-450 CYP1A2**
4. **C0608437|CYP2C9 protein, human**
5. **C0960580|CYP2C19 protein, human**
6. C1412027|Xenobiotic Monooxygenases

3. Conclusions

Some concepts contain strings which should be marked as suppressible. Many of these strings are already marked suppressible for a given UMLS release; this report recommends further cases some of which are universally applicable and some of which are appropriate in more limited environments such as the natural language processing done by MetaMap.

The analysis in this and previous editions of this report reveals some classes of ambiguity commonly occurring in the Metathesaurus:

- **Contextual (or hierarchical) ambiguity.** This class of false ambiguity is exemplified by the string 'Prostate' for 'Prostatic Diseases'. It normally arises from terms which require context within their vocabulary (in this case, a disease hierarchy) in order to be properly understood. Contextual ambiguities can be classified according to their participants:

- **Body part/disease ambiguity** exemplified by 'Prostate' and 'Prostatic Diseases'
- **Body part/procedure ambiguity** exemplified by 'Stomach' and 'Procedures on the stomach'
- **Pathology/procedure ambiguity** exemplified by 'Pathology' and 'Pathology procedure'
- **Medical device/procedure ambiguity** exemplified by 'Prosthesis' and 'Prosthesis Implantation'
- **Substance/therapy ambiguity** exemplified by 'Anthracyclines' and 'prior anthracycline therapy'
- **Substance/measurement ambiguity** exemplified by 'Thyroid stimulating immunoglobulins' and 'Thyroid stimulating immunoglobulins assay'

- **Generalization ambiguity.** This is also false ambiguity caused by grouping several concepts together using a more general term. For example, 22 concepts including 'Protocols: Activities' and 'Protocols: Pre- or Intra- or Post-Procedure' are generalized to 'Protocols' which does seem to be a legitimate synonym of the concept 'Protocols documentation'.

- **Meta ambiguity.** This new class of ambiguity, represented by strings such as 'Other complications of procedures NEC in ICD10', contain meta information. In this case it is the name of the vocabulary, ICD10 in the example. As opposed to the first class of ambiguity above in which strings such as 'Prostate' meaning 'Prostatic Diseases' do not say enough about themselves, these strings say too much. It is true that the meaning of a string containing 'NEC', 'not elsewhere classified' or like phrase, depends upon its vocabulary, but such information is already available in the MSRO file (where it belongs). It is also true that such strings have different meanings and strictly speaking should be different concepts. But the practical result of such a representational scheme is to introduce an ambiguity that most users do not want or need to resolve. (It is not even clear that those who might want to resolve the ambiguity can do so with the information available in the Metathesaurus.)

- **Abbreviation ambiguity.** This is another, large class of ambiguity caused by distinct concepts having the same acronyms (or abbreviations). An example from above is that 'Mitral Valve Stenosis', 'Multiple Sclerosis', 'Morphine Sulfate' and 'millisecond' all have abbreviation 'MS' or 'ms'. Although this class represents true ambiguity in a strict sense, it is better to disallow it in many text processing situations, especially those in which authors define the abbreviations they use. Unlike the other classes of ambiguity defined above, we do not recommend that this case be reflected in changes to the Metathesaurus. This kind of ambiguity will be suppressed for MetaMap processing only.