MATERIAL Query Overview

Center for Advanced Study of Language (CASL)
October 2017
• Three types of queries
  – simple
  – conceptual
    • full conceptual
    • EXAMPLE_OF
  – hybrid

• Query constraints
  – semantic
  – morphological
Query type I: simple queries

• Search goal:
  – find documents containing a translation equivalent of the query expression

• Relevant documents contain the English query string or a word or phrase that could be reasonably translated as the query string in the given context
Simple queries: some examples (1/2)

• query string: campaign
  – relevant: “campaigns”, “campaigning”

• query string: "pet cat"
  – relevant: “my pet cat”, “pet Persian cat”
  – not relevant: “pet Persian”
Simple queries: some examples (2/2)

• query string: soap
  – relevant: “a bar of soap”, “soap dispenser”, not relevant: “detergent”

• query string: brother, auction
  – relevant: “the brothers auctioned off their record collection”, “Auctions are fun. ... My brother works with prisoners.”
  – not relevant: “the brothers sold their record collection to the highest bidder”
Query type IIa: full conceptual queries

• Search goal:
  – Find documents with content relevant to the topic introduced by the query term
• Relevant documents mention the topic or a subtype or instance of the topic, or touch on all the distinctive features of the topic
Conceptual queries: some examples

• query string: **boat**+
  – relevant: “catamaran”, “dinghy”
  – not relevant: “ocean”, “floating”

• query string: **safari**+
  – relevant: “guided tours of African wildlife refuges”, “big game hunting expeditions in East Africa”
  – not relevant: “a cruise along the Nile”, “wine-tasting tour in South Africa”
Query type IIb: EXAMPLE_OF queries

• Search goal:
  – Find documents mentioning an example of the query term

• Relevant documents mention a subtype or instance of the topic
EXAMPLE_OF queries: some examples

• query string: EXAMPLE_OF(virus)
  – relevant: “HIV”, “influenza”
  – not relevant: “disease”

• query string: EXAMPLE_OF(head of government)
  – relevant: “the Prime Minister”, “Governor”
  – not relevant: “Speaker of the House”
Query type III: hybrid queries

• Logical conjunction of a simple query and either a full conceptual or EXAMPLE_OF query

• Relevant documents satisfy the conditions of both elements of the query

• The conditions of the two elements of the query can be satisfied by two separate parts of a document
Hybrid queries: an example

• query string: fruit+, "migrant worker"
  – relevant: “few migrant workers could be found to harvest strawberries this year”
  – not relevant: “in past years, fruit farmers hired large numbers of Guatemalan laborers”
  – not relevant: “Government reduces quota on migrant workers”
Semantic query constraints

- Used to clarify the intent of a query string
- Three types
  - Synonym
  - Hyponym
  - Event frame
Semantic query constraints

• **Synonym**: roughly the same meaning as query string
  – fly [syn: to pilot]

• **Hypernym**: concept that encompasses but is more general than the query string
  – fly [hyp: insect]

• **Event frame**: context in which the query string has a particular meaning
  – fly [evf: baseball]
Morphological constraints

• To be relevant, a word or phrase must match the “marked” morphological properties of the query string (e.g. past tense on verbs; plural on nouns)

• Only simple queries can be morphologically constrained
Morphological constraints: examples

<contaminated>: in order to be relevant, a word or phrase must be equivalent to the English past tense or past participle

<airplanes>: in order to be relevant, a word or phrase must be equivalent to the English plural