Challenges of Natural Language

Semantics (vocabulary)

- What do the words *mean*?
- Meaning may be context-sensitive.
  > "The bandage was wound around the wound."
- A single word (i.e., sequence of characters) may have *many* meanings.

Syntax (grammar)

- Rules may be "flexible" and poorly understood.
- The language is NOT commutative.
  > "Santa comes only at Christmas." vs "Only Santa comes at Christmas."

Idioms

- Common usage may be counter-intuitive (and even strictly incorrect).
- "Just tell him I sent you, and Bob's your uncle."
- "bite the dust", "say uncle", "a poor man's _______", "got my goat"

Challenges of Natural Language

Spelling – don't even get me started on that one…

- "ghoti" may sensibly be pronounced "fish"
- No real effort to standardize until the late 18-th century.
- Few rules, and those are confusing (if not confused).
- A mis-spelled word may well be another, valid word.

Eye halve a spelling chequer
It came with my pea sea,
It plainly marques four my revue
Miss steaks eye kin knot sea.
Eye strike a key and type a word
And weight for it two say,
Weather eye and wring oar write
It shows me strait a weigh.
As soon as a mist ache is maid
It nose bee fore two long,
And eye can put the error rite
Its rare lea ever wrong.
To rite with care is quite a feet
Of witch won should bee proud,
And wee mussed dew the best wee can,
Sew flaw's are knot aloud.
Eye have run this poem threw it
Your sure reel glad two no,
Its letter perfect awl the weigh
My chequer tolled me sew.
-Sauce unknown

Challenges of Natural Language

Ambiguity --- the condition that information can be *fairly* interpreted or understood in more than one way.

He ate the cookies on the couch.

You can't put too much water on a nuclear reactor core.
(http://www.youtube.com/watch?v=6JrIYR8jArk)

Some headlines:

- Iraqi head seeks arms (semantic ambiguity, doubly so)
- Teacher strikes idle kids (part-of-speech ambiguity)
- Stolen painting found by tree (structural ambiguity)

The Many Forms of If…Then

In English, there are many different ways to say that "P implies Q", including:

If rain is falling, then the sky is cloudy.

Rain is falling only if the sky is cloudy.

The sky is cloudy, if rain is falling.

The sky is cloudy, provided that rain is falling.

The sky is cloudy, assuming that rain is falling.

In order for the sky to be cloudy, it is sufficient that rain be falling.

In order for rain to be falling, it is necessary that the sky be cloudy.

Negation

Forming the correct negation of a statement in English seems to be particularly problematic for many people:

> If rain is falling, then the sky is cloudy.

The negation of a statement must have the property that it has the opposite truth value from the original statement*.

Recalling the analysis of the truth values of "if P then Q", we see that the negation must *affirm the antecedent* (P) *and deny the consequent* (Q).

So, the negation of the statement above is

> Rain is falling, and the sky is not cloudy.

*** But, simply having the opposite truth value isn't sufficient.**

Negation

Try negating the following:

Alfred will pass, if he concentrates.

Alfred concentrated, but he did not pass.

If logic is not difficult, Alfred will pass.

Logic is not difficult, but Alfred did not pass.

Negation with AND

What is the negation of:

> Helium is denoted by He and hydrogen is denoted by Hy.

An AND statement is true if and only if both parts are true. So, an AND statement is false if and only if at least one part is false.

So, the negation of the statement above would be:

> Helium is not denoted by He or hydrogen is not denoted by Hy.

Negation with OR

What is the negation of:

Tin is denoted by Tn or iron is denoted by Fe.

An OR statement is true if and only if at least one part is true. So, an OR statement is false if and only if both parts are false.

So, the negation of the statement above would be:

Tin is not denoted by Tn and iron is not denoted by Fe.

Formal Negation Rules

The negation of an if… then is given by:

"not (if P then Q)" is equivalent to "P and (not Q)"

DeMorgan's Laws state that:

"not (P and Q)" is equivalent to "(not P) or (not Q)"

"not (P or Q)" is equivalent to "(not P) and (not Q)"

These should be intuitively obvious, but people commonly get this wrong when dealing with statements expressed in natural language.

Quantification

Quantifier --- a type of determiner, such as all or many, that indicates quantity.

All men are mortal.

Every prime integer larger than 2 is odd.

Some headlines exhibit ambiguity.

Most mammals are terrestrial rather than aquatic.

A few birds are naturally flightless.

Quantification: Order Matters

Every question on the WASI was answered correctly by some student in the class.

Some student in the class correctly answered every question on the WASI.

For every question on the WASI, there exists a student in the class who answered that question correctly.

There exists a student in the class who answered every question on the WASI correctly.

Quantification: Domains

When a quantifier is used, there should usually be an associated *domain*; that is, we should know the collection of "things" that are the object of the quantifier.

Some student correctly answered every question ~~on the WASI~~.

In the absence of a specific domain, the sentence may be ambiguous or meaningless.

Every x is prime.

Quantification: Negation

What is the negation of:

Every prime integer is odd.

Clearly, the negation must deny the assertion that every prime integer has the stated property.

That means the negation must merely claim that there is some prime integer that does not have the stated property:

There is a prime integer that is not odd.

Quantification: Negation

What is the negation of:

Some three-digit integer equals the product of its digits.

Clearly, the negation must deny the assertion that there is an integer with three digits has the stated property.

That means the negation must merely claim that there is no integer with three digits that does have the stated property:

No three-digit integer equals the product of its digits.

Alternatively:

Every three-digit integer is unequal to the product of its digits.

Quantification: Negation and the Domain

Note that negating does not alter the domain:

All dogs are endothermic.

Some dogs are not endothermic.

You cannot deny the truth of the first statement by talking about the wrong set of things:

Some iguanas are not endothermic.

Quantification: Formal Negation

So:

"not (every x in D has property P)" is
"some x in D does not have property P"

and

"not (some x in D has property P)" is
"no x in D has property P"

Of course, it gets more interesting; negate these:

For every natural number N, there is a natural number M such that $M < N$.

There is a natural number M such that, for every natural number N, $M \leq N$.

For every integer N, if N is a multiple of 2 then $\log_2(N)$ is an integer.

For every real number $\varepsilon > 0$, there is a real number $N > 0$ such that whenever x is a real number such that $x > N$, $1/x < \varepsilon$.

Answers

The negations:

There is a natural number N such that for every natural number M, M < N.

For every natural number M, there is a natural number N such that M > N.

There is an integer N such that N is a multiple of 2 and $\log_2(N)$ is not an integer.

There is a real number ε > 0 such that, for every real number N > 0, there is a real number x such that x > N, but 1/x ≥ ε.

Six Myths about Reading

What speed-reading advocates say:

1. Don’t subvocalize when you read
2. Read only the key words
3. Don’t be a word-by-word reader
4. Read in thought groups
5. You can read at speeds of 1000 or more words a minute – without any loss of comprehension
6. Don’t regress or re-read

These suggestions are simply incorrect, in almost all cases.

There are no short cuts to comprehension!

Difficult material typically requires re-reading to understand.

Whimbey, Ch V