

Natural Language Math Problem Solving

Stanford Lockhart

Niclas Skaalum

Geoff Caven

Chelsey Childs

Math as a Natural Language

— — —

Mathematical expressions can be expressed in natural language.

Eg: “ $2 + 2$ ” == “two plus two” == “sum of two and two” == ...

Naturally expressed mathematics uses a fairly limited dictionary.

NLTK

Toolkit in python for processing human languages

Provides tools to stem, tokenize, tag, classify, create grammars

`nltk.PCFG`

`nltk.stem WordNetLemmatizer`

`nltk.parse pchart`

Grammar

```
S -> S ADD S
S -> ADD S 'and' S
S -> 'subtract' S
S -> THOU
S -> T
T -> T MUL T
T -> MUL T 'and' T
T -> 'subtract' T
T -> THOU
ADD -> 'add' | 'subtract'
MUL -> 'multiply' | 'divide'
THOU -> HUN
THOU -> CD 'thousand' HUN
THOU -> CD 'thousand' 'and' HUN
THOU -> CD 'thousand'
```

```
HUN -> TEEN
HUN -> CD 'hundred' TEEN
HUN -> CD 'hundred' 'and' TEEN
HUN -> CD 'hundred'
TEEN -> CD
TEEN -> TEN CD
TEEN -> TEN
TEEN -> 'ten'
TEEN -> 'eleven'
...
TEN -> 'twenty'
TEN -> 'thirty'
...
CD -> 'zero'
CD -> 'one'
...
```

Sources

— — —

<http://www.abstractmath.org/Handbook/handbook.pdf>

Wells, C. (2003, March). A handbook of mathematical discourse. PA: Infinity.

<http://annualreviews.org/doi/pdf/10.1146/annurev-linguist-030514-125312>

Liang, P., & Potts, C. (2015). Bringing machine learning and compositional semantics together. *Annu. Rev. Linguist.*, 1(1), 355–376.

Ambiguities

— — —

Negatives

minus two minus minus two

$-(2 - (-2))$ or $(-2) - (-2)$

BEDMAS

eight plus two times four

$8 + (2 * 4)$ or $(8 + 2) * 4$

Parsing

Tree is built by NLTK

Calculate tree bottom up

BEDMAS means multiplication and division are at the bottom of the tree

Value is returned by the root node

Example

-- --

minus two minus minus two

$$((-2)-(-2)) = 0$$

```
(S
  (S subtract (S (THOU (HUN (TEEN (CD two))))))
  (ADD subtract)
  (S subtract (S (THOU (HUN (TEEN (CD two))))))
)
```


Example

— — —

product of one hundred and three
and four

$$(103 \times 4) = 412$$

```
(S
  (T
    (MUL multiply)
    (T (THOU (HUN (CD one) hundred and (TEEN (CD three))))
    and
    (T (THOU (HUN (TEEN (CD four))))
  )
)
```

Example

— — —

eight plus two times four

$$(8+(2*4)) = 16$$

```
(S
  (S (THOU (HUN (TEEN (CD eight))))))
  (ADD add)
(S
  (T
    (T (THOU (HUN (TEEN (CD two))))))
    (MUL multiply)
    (T (THOU (HUN (TEEN (CD four))))))
  )
)
)
```