



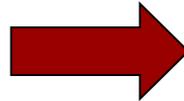
# English to Olog Translation

or: how I learned to stop worrying and love the Olog.

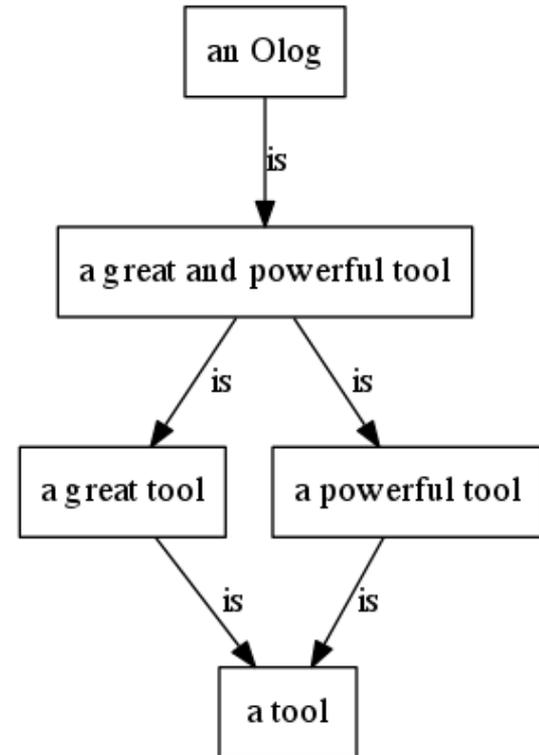
English

An Olog is a great and powerful tool.

Magic?



Olog Olog

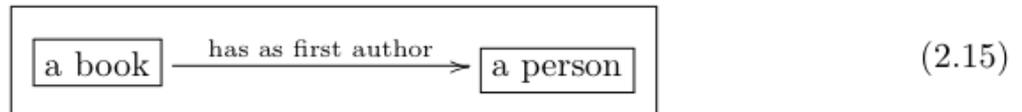
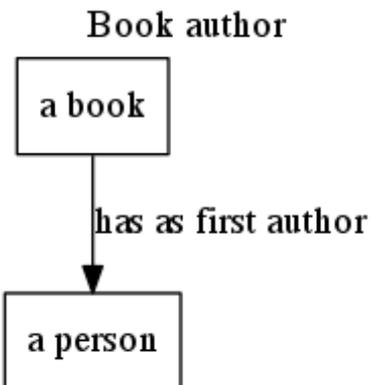


A proof of concept  
by a M.I.T. Student

# Inspiration

## 2.3.2.4 Reading aspects and paths as English phrases

Each arrow (aspect)  $X \xrightarrow{f} Y$  can be read by first reading the label on its source box (domain of definition)  $X$ , then the label on the arrow  $f$ , and finally the label on its target box (set of values)  $Y$ . For example, the arrow



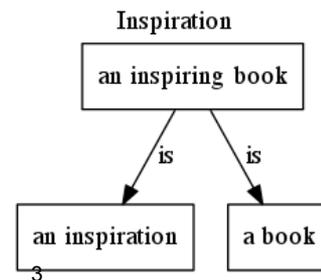
is read “a book has as first author a person”.

*Rules of good practice 2.3.2.8.* An aspect is presented as a labeled arrow, pointing from a source box to a target box. The arrow text should

- (i) begin with a verb;
- (ii) yield an English sentence, when the source-box text followed by the arrow text followed by the target-box text is read; and
- (iii) refer to a functional relationship: each instance of the source type should give rise to a specific instance of the target type.

# Inspiration continued

- If aspects can be read as English sentences... can English sentences be turned into aspects?
- Could we take a list of aspects written in English, and back out the Olog?
  - Perhaps even more inferences could be made
- **DISCLAIMER** – Because English is not always functional... we will not always get true Ologs out of our sentences...
  - Monads are the solution



# Kleisli Category – relations are cool

- Instead of  $X \rightarrow Y$ 
  - Where the arrows in the Olog represent functions
- We can use  $X \rightarrow \mathbf{P}Y$ 
  - Where the arrows in this context represent relations
  - The powerset of  $Y$  contains all possible subsets of  $Y$ , and so any relation from  $X$  to  $Y$  is kosher.
- **This allows arrows in the Olog to represent relations instead of functions.**

# Mission Statement and Approach

## ➤ Mission Statement

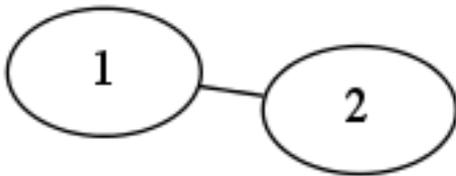
- Create a proof of concept that can take a list of English sentences, pull out some relationships, and display them as an Olog.
- People speak English... why can't we translate for them?

## ➤ Decompose the problem

- Convert English sentences to parts of speech
- Given parts of speech, find boxes and arrows
- Given boxes and arrows, display the resulting Olog automatically

# Draw Ologs Automatically

- Searched for a tool for drawing graphs
  - Found a program called Graphviz
    - “Graph Visualization Software” developed by AT&T
    - Uses DOT language scripts.
    - It’s free
  - The kicker... it has a Python wrapper called PyGraphviz!
    - `A.add_edge(1,2)`



# English sentences to parts of speech

➤ Searched for a tool for identifying parts of speech (POS) in a sentence

➤ Found a python module called `topia.termextract`

➤ This module can tokenize a sentence...

```
>>> tagger.tokenize('This is a simple example.')
```

```
['This', 'is', 'a', 'simple', 'example', '.']
```

➤ And then determine the parts of speech

```
>>> tagger('This is a simple example.')
```

```
[['This', 'DT', 'This'],
```

```
['is', 'VBZ', 'is'],
```

```
['a', 'DT', 'a'],
```

```
['simple', 'JJ', 'simple'],
```

```
['example', 'NN', 'example'],
```

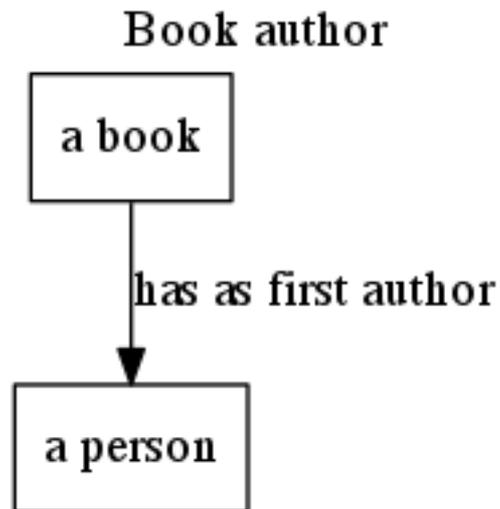
```
['.', '.', '.']]
```

# Given POS, find boxes and arrows

- With PyGraphviz and topia.termextract both available for free, coding the POS -> (boxes and arrows) portion in Python made the most sense
- After looking at a few Ologs, a method for doing this was determined (assuming a NF, VF, NF structure):
  1. First NF starts at the beginning and goes until the first verb.
  2. VF starts at the verb and goes until the first “a” or “an”.
  3. And the second NF goes from the “a” or “an” to the end.

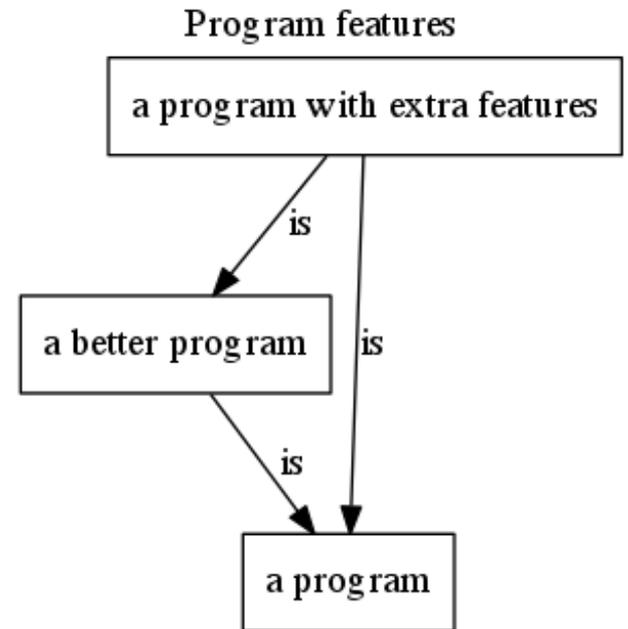
# A simple example

- Book author
- A book has as first author a person.



# More features

- Recursively tear down adjective phrases
- Deal with prepositional phrases
- Remove adverbs from verb phrases
- Check box for verbose output
- Open Source
- Decided to avoid pull outs for now
  - Blows up quickly



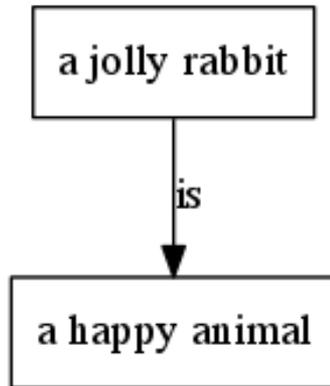
# Adjective phrases

## Standard Output

### Adjective phrases

A jolly rabbit is a happy animal.

### Adjective phrases

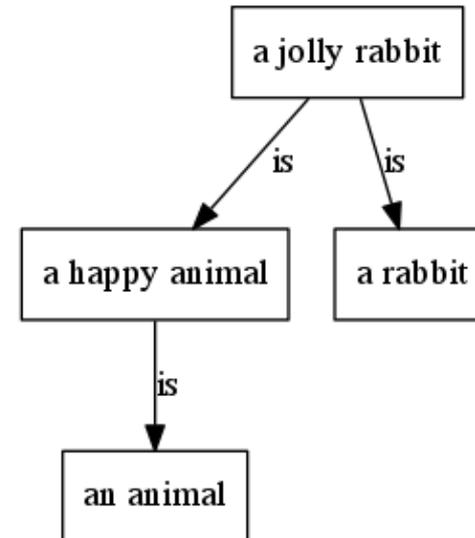


## Verbose Output

### Adjective phrases

A jolly rabbit is a happy animal.

### Adjective phrases

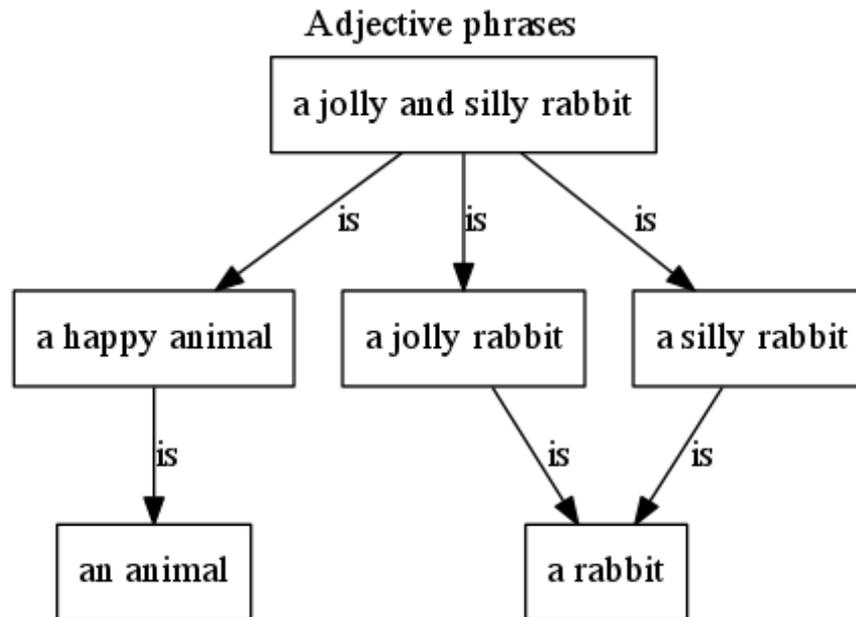


# More Adjective Phrases

➔ What happens if we add more adjectives?

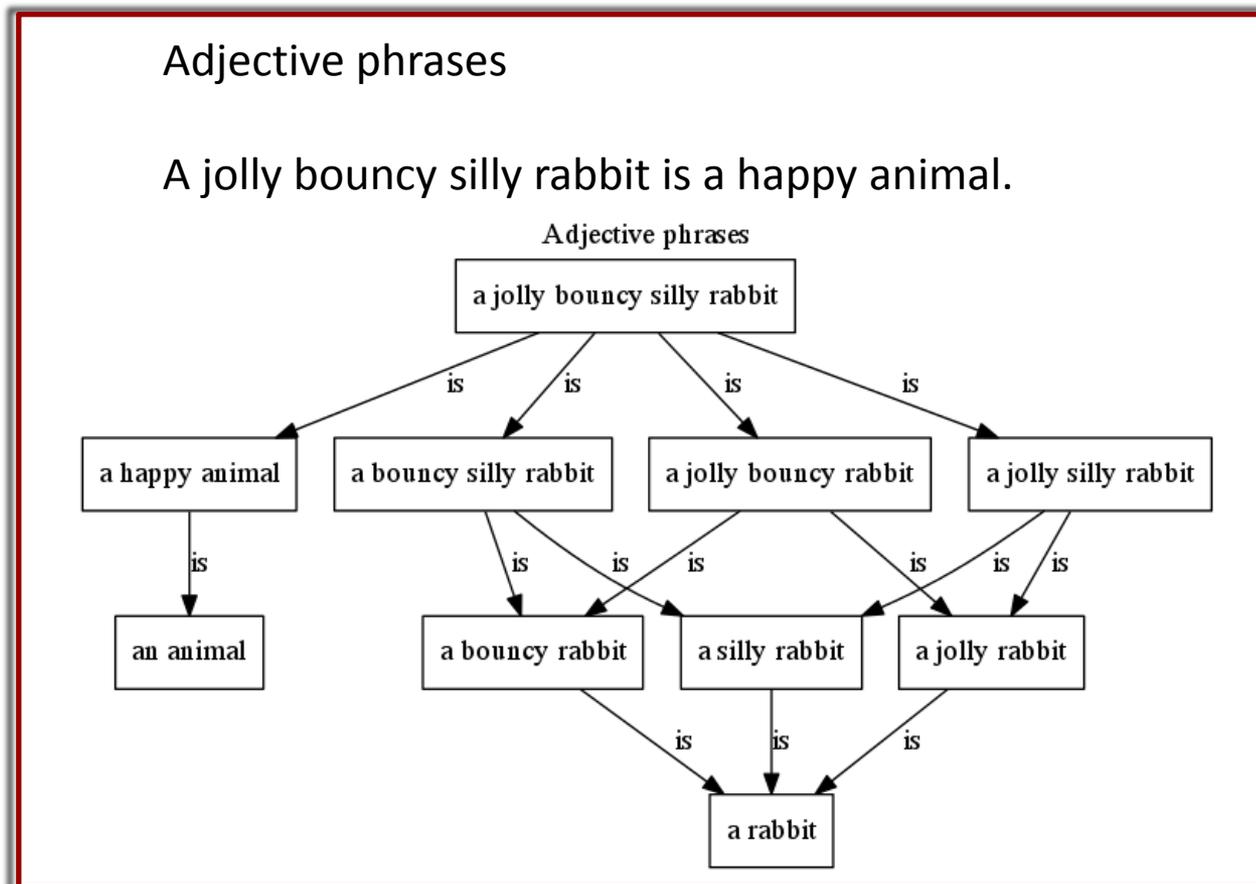
Adjective phrases

A jolly and silly rabbit is a happy animal.



# Even More Adjective Phrases

➔ What happens if we add even more adjectives?





# Prepositional phrases

## Standard Output

Prepositional phrases

A man with a broom is a person.

Prepositional phrases

a man with a broom

is

a person

## Verbose Output

Prepositional phrases

A man with a broom is a person.

Prepositional phrases

a man with a broom

is

is

a person

a man

# More Prepositional Phrases

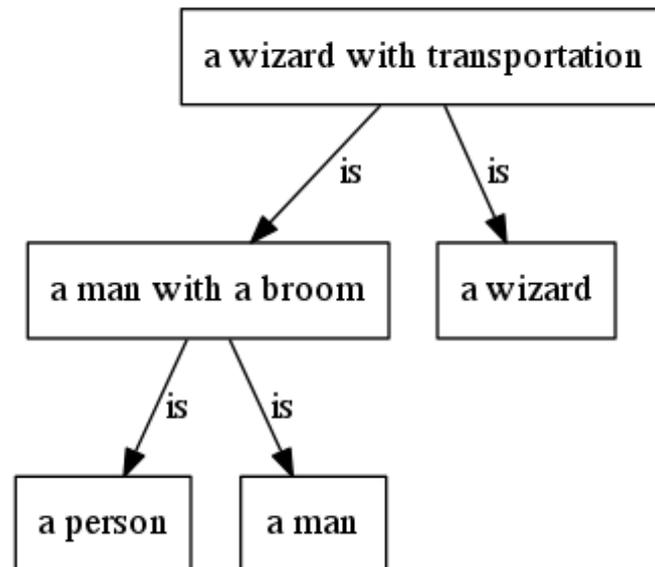
➔ What happens if we add more sentences?

Prepositional phrases

A man with a broom is a person.

A wizard with transportation is a man with a broom.

Prepositional phrases



# Even More Prepositional Phrases

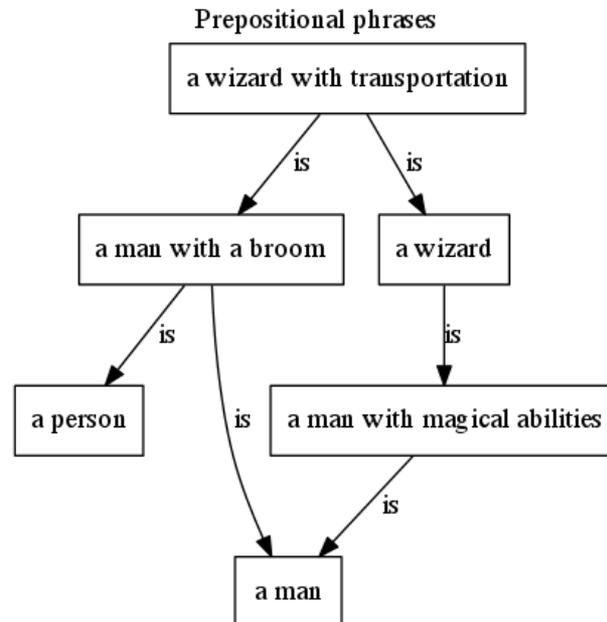
➔ What happens if we add even more sentences?

Prepositional phrases

A man with a broom is a person.

A wizard with transportation is a man with a broom.

A wizard is a man with magical abilities.



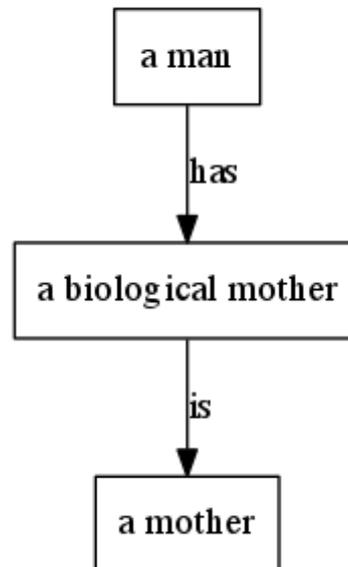
# Adverbs

➔ We can remove adverbs from verb phrases...

Remove adverbs

A man always has a biological mother.

Remove adverbs



# More examples

This is a test

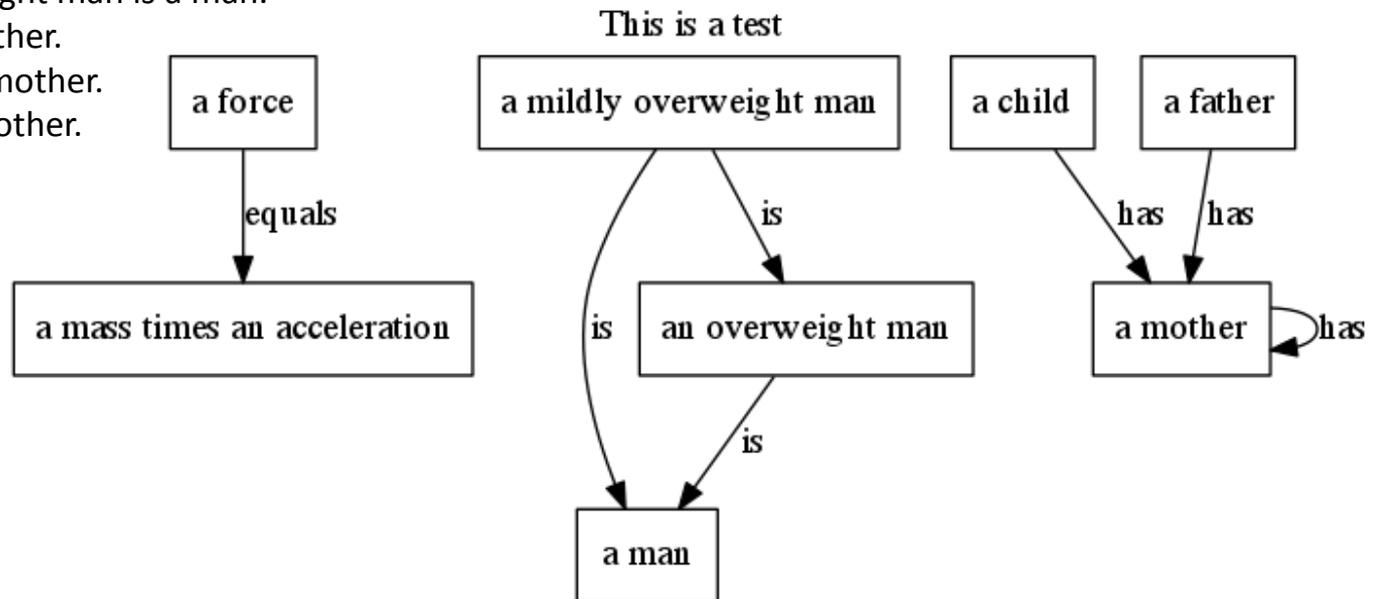
A force equals a mass times an acceleration.

A mildly overweight man is a man.

A child has a mother.

A mother has a mother.

A father has a mother.





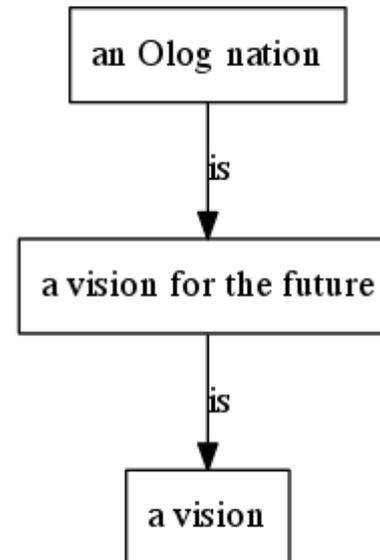
# A vision for the future

- Additional Features
  - Web/graphical interface
    - Like Google translate
  - Noun phrase parsing
  - Olog to English output
  - FQL output
  - Proper nouns to data
- Future applications
  - Olog wikipedia and the dictionary
  - Scientific paper Ologging

Vision for the future.

An Olog nation is a vision for the future.

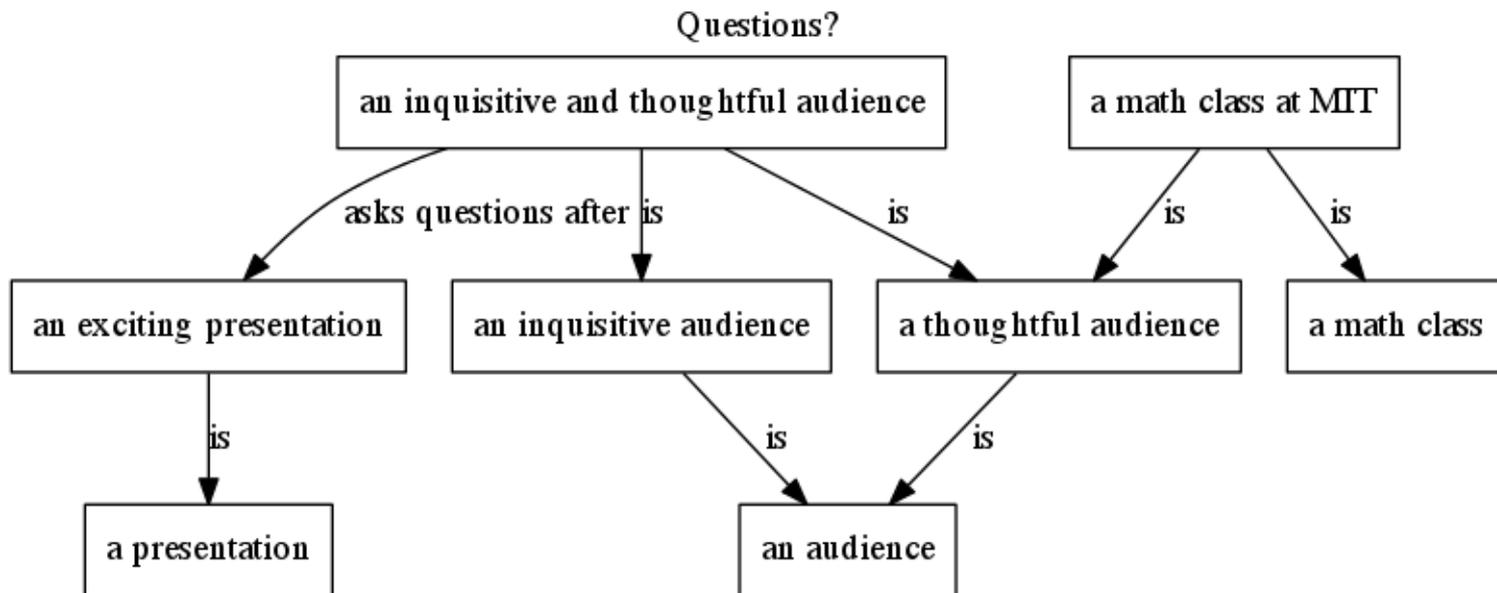
Vision for the future



# Questions and comments?

## ➤ Questions?

- An inquisitive and thoughtful audience asks questions after an exciting presentation.
- A math class at MIT is a thoughtful audience.



<http://sourceforge.net/projects/olognation/files/English2Olog.py/download>

MIT OpenCourseWare  
<http://ocw.mit.edu>

18.S996 Category Theory for Scientist  
Spring 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.