UNIVERSITY OF CAMBRIDGE

# Applying Language Models to Language Learning

Duncan Roberts

MPhil in Advanced Computer Science

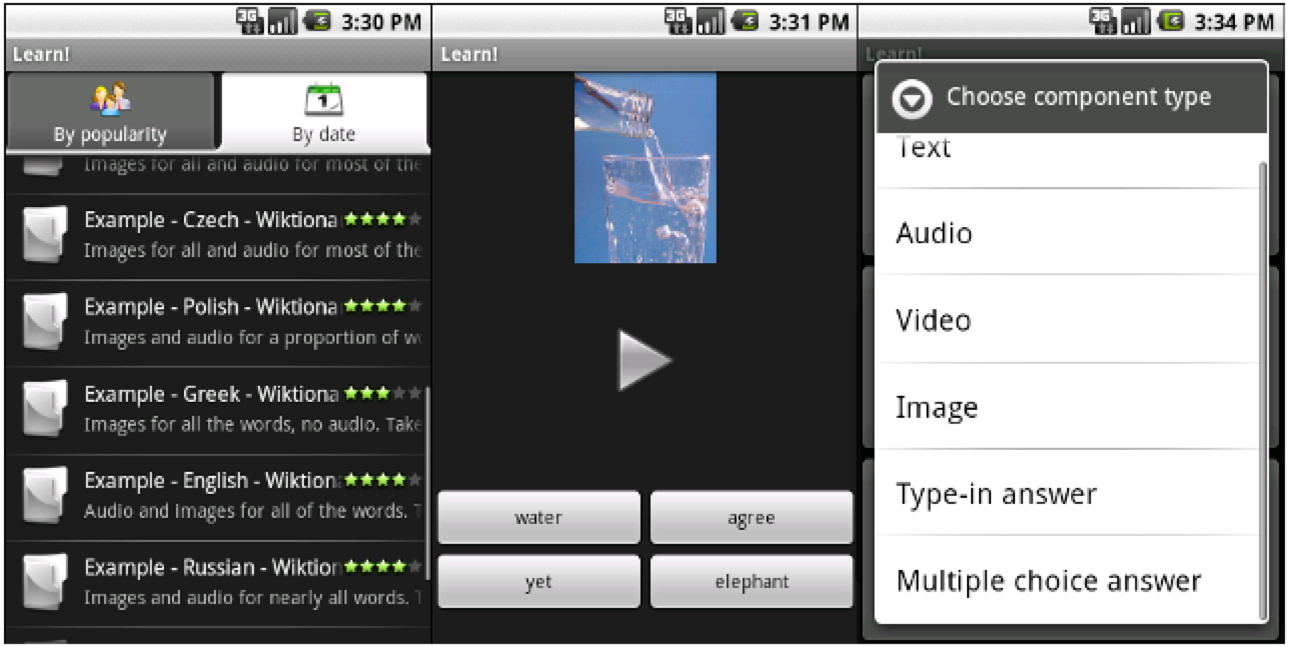
# Digital flash cards!

Popular approach to learning. DTG's own *Learn!* Android app has sets for:

- Basic chemistry
- US Navy core values
- Career development strategies
- ...Languages?

# Language learning through flash cards

Teaching vocab not too hard:



Images for all and audio for most of the

Example - Czech - Wiktiona ★★★★
Images for all and audio for most of the

Example - Polish - Wiktiona ★★★★
Images and audio for a proportion of w

Example - Greek - Wiktiona ★★★
Images for all the words, no audio. Take

Example - English - Wiktion ★★★★
Audio and Images for all of the words. T

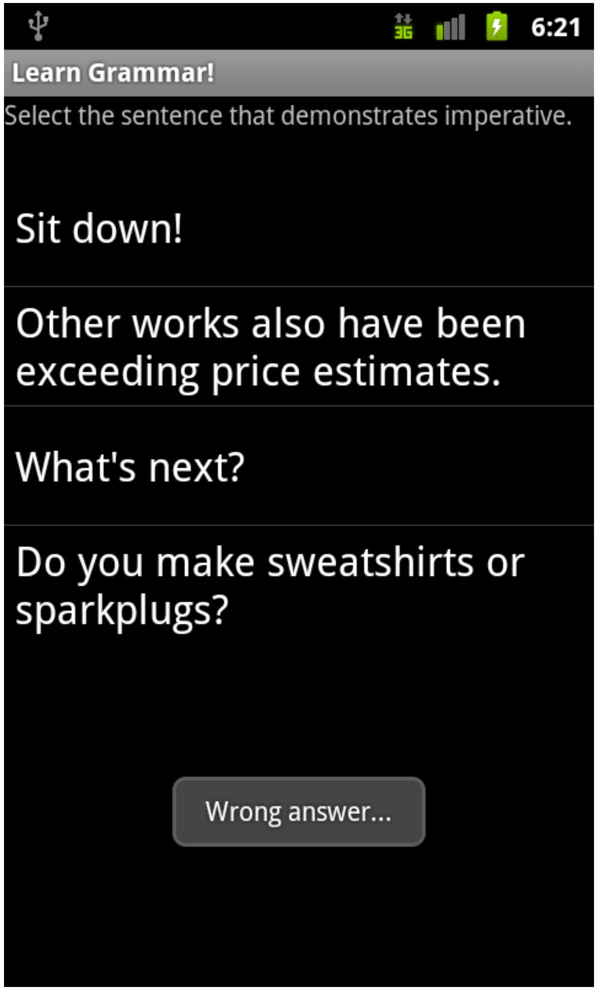
Example - Russian - Wiktion ★★★★
Images and audio for nearly all words. T

# Language learning through flash cards

But grammar?

# Teaching grammar

- Need to make it quick and easy to create large sets of examples and counter-examples of arbitrary grammar points, e.g.
  - Future tense
  - Relative clauses
  - The imperative
  - Pretty much anything else – code should be task-independent



Select the sentence that demonstrates imperative.

Sit down!

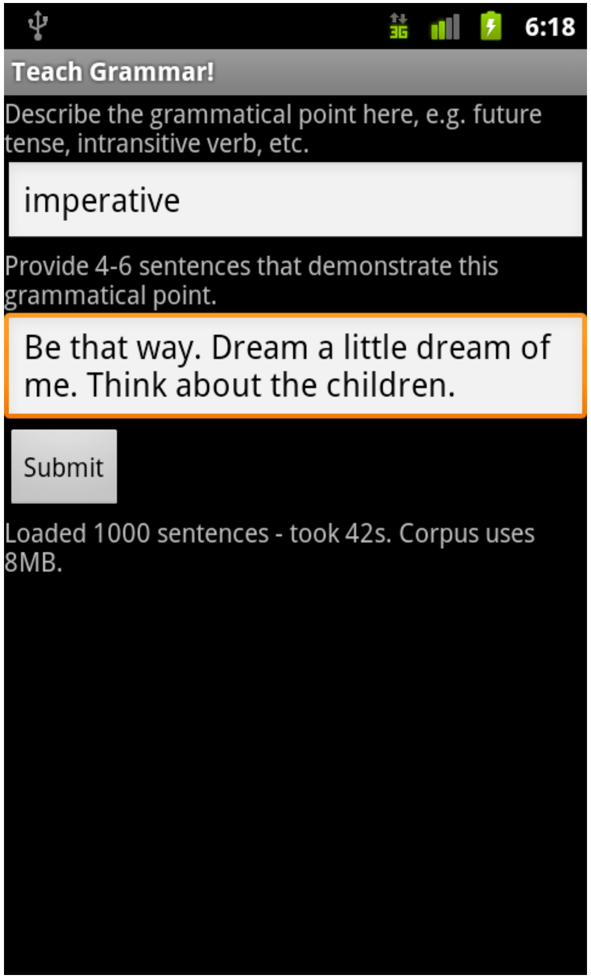
Other works also have been exceeding price estimates.

What's next?

Do you make sweatshirts or sparkplugs?

# Approach

1. Prompt 'teacher' for:
   - A title for our learning task (uninterpreted)
   - ~5 example sentences
   - 5 counter-example sentences
2. Work out what examples have in common (that's not in a counter-example)
3. Find many other sentences in a pre-parsed corpus exhibiting this *commonality*



Describe the grammatical point here, e.g. future tense, intransitive verb, etc.

imperative

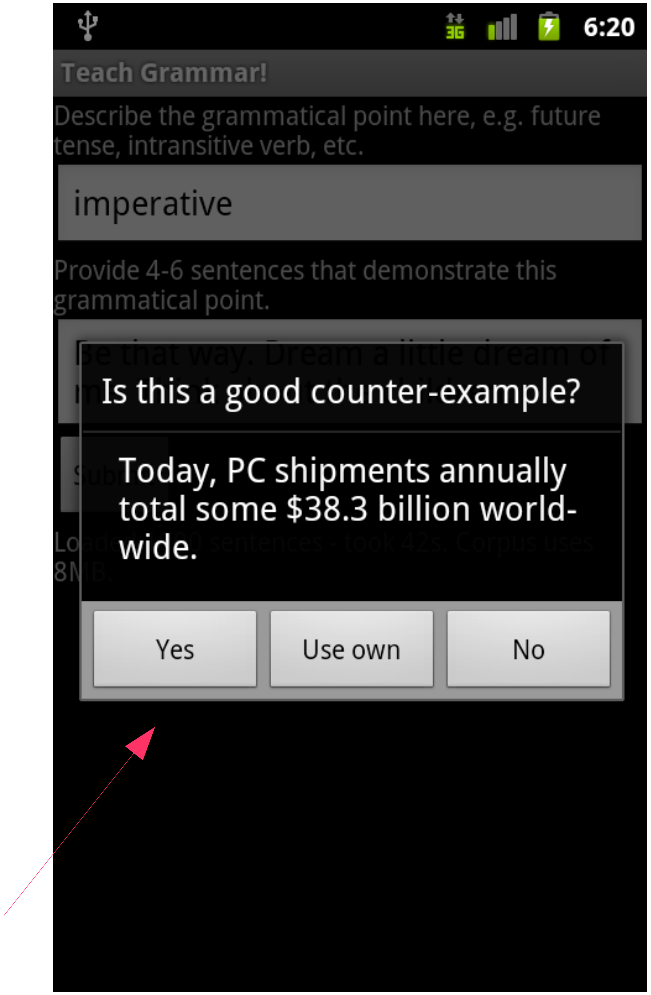
Provide 4-6 sentences that demonstrate this grammatical point.

Be that way. Dream a little dream of me. Think about the children.

Loaded 1000 sentences - took 42s. Corpus uses 8MB.

# Approach

1. Prompt 'teacher' for:
   - A title for our learning task (uninterpreted)
   - ~5 example sentences
   - 5 counter-example sentences
2. Work out what examples have in common (that's not in a counter-example)
3. Find many other sentences in a pre-parsed corpus exhibiting this *commonality*



Describe the grammatical point here, e.g. future tense, intransitive verb, etc.

imperative

Provide 4-6 sentences that demonstrate this grammatical point.

Be that way. Dream a little dream of

Is this a good counter-example?

Today, PC shipments annually total some $38.3 billion world-wide.

*Save user typing by suggesting random sentences as counter-examples*

# Analysing a sentence

- Each word has*:
  - A part of speech (POS) tag: plural noun, adjective, past tense verb...
  - A base form ('lemma'): *thought* → *think*

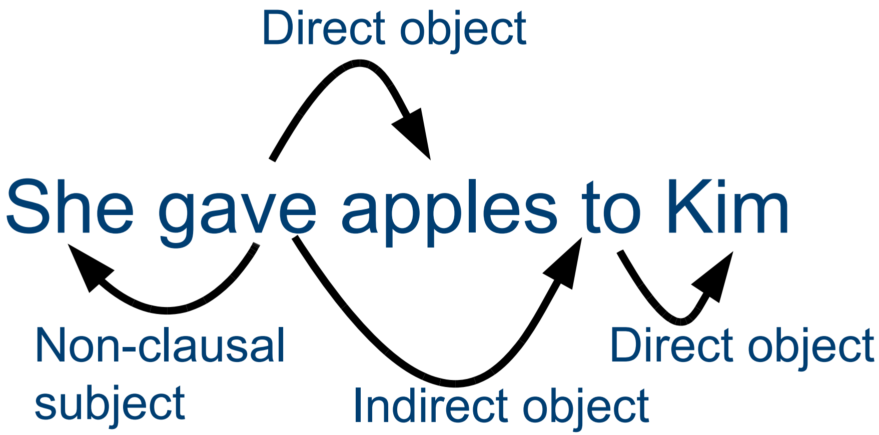
*Most features derived using C&C – Stephen Clark's syntactic parser

# Analysing a sentence

| Word | Lemma | POS |
|---|---|---|
| She | she | Personal pronoun |
| gave | *give* | Past tense |
| apples | *apple* | Plural common noun |
| to | to | Preposition |
| Kim | Kim | Proper noun |

# Analysing a sentence

We can also extract *grammatical relations* (GRs) between words:



Direct object

She gave apples to Kim

Non-clausal subject

Indirect object

Direct object

# Commonality

What do the example sentences **have in common**?

- Used to find other, similar sentences

# Commonality – case study #1

Consider the “going to” future tense, e.g.

- He's going to be rich one day.
- I'm going to think about it.
- Becky is going to read her book.

(But *not “I'm going to the shop to buy some tea” - present continuous*)

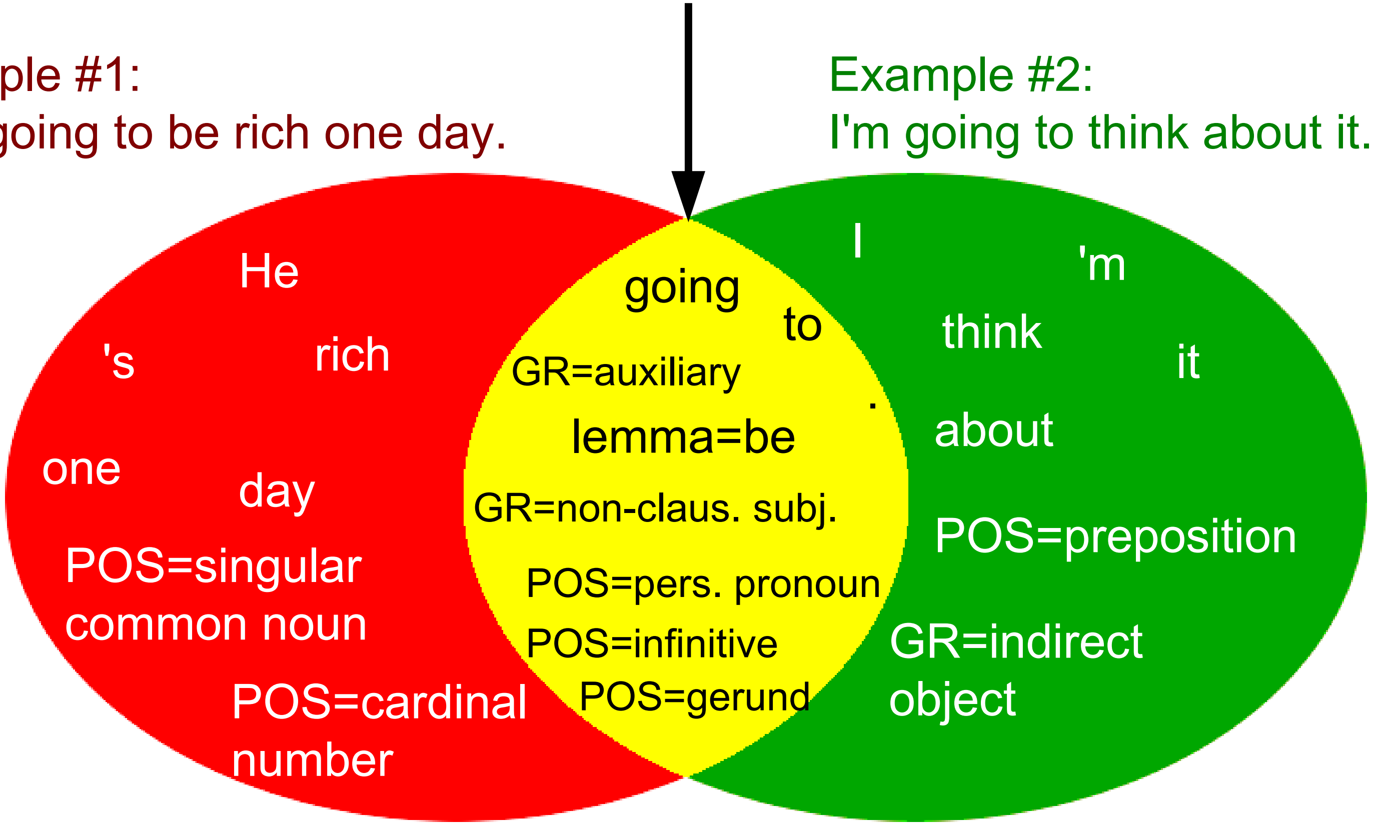
Can't spot imperative using approaches described so far...

# Weak commonality

*Weak* commonality: **intersection** of features observed in examples

Example #1:
He's going to be rich one day.

Example #2:
I'm going to think about it.

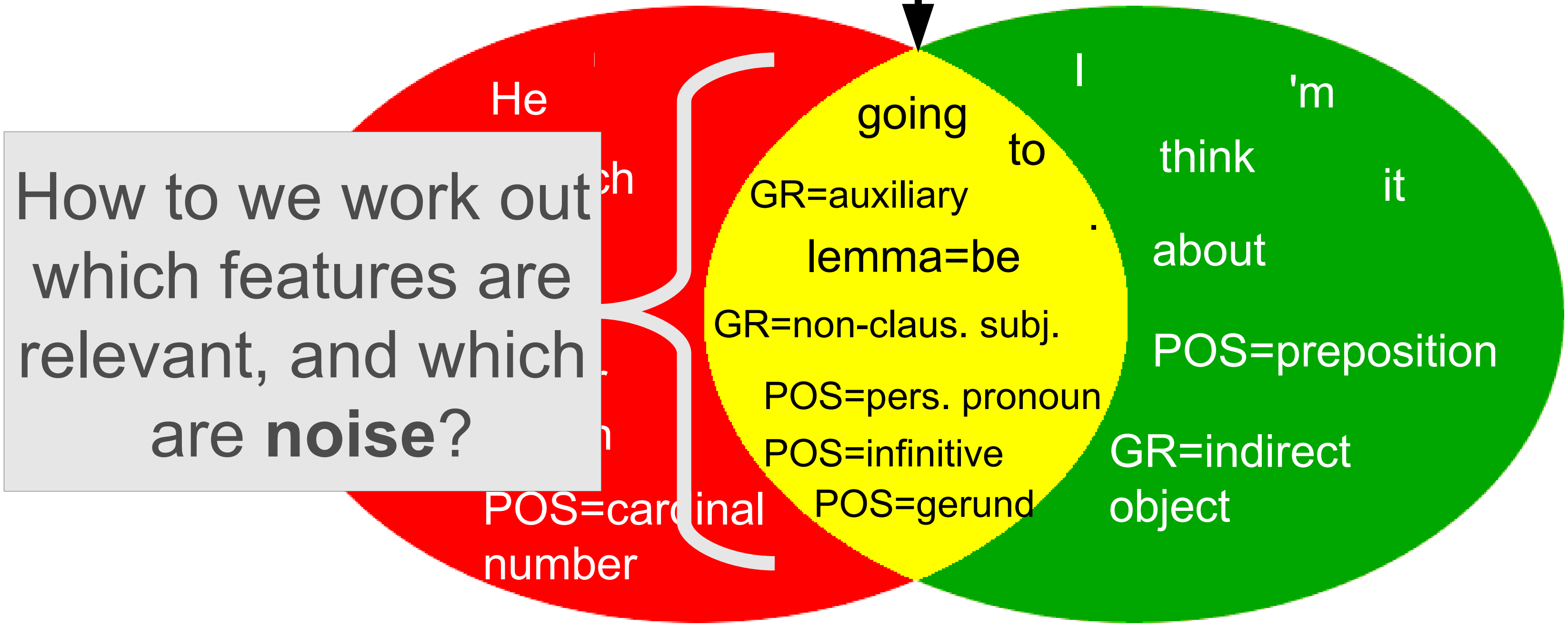


# Weak commonality

*Weak* commonality: **intersection** of features observed in examples

Example #1:
He's going to be rich one day.

Example #2:
I'm going to think about it.



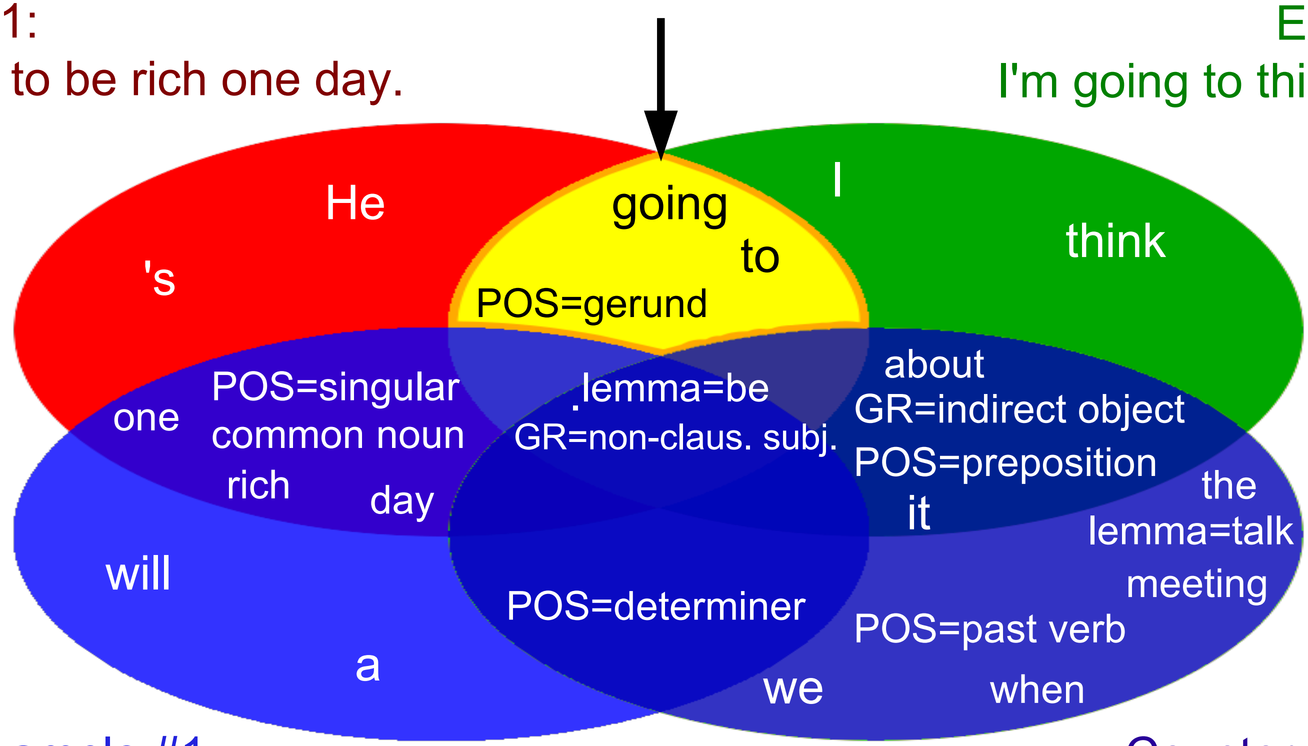
How to we work out which features are relevant, and which are **noise**?

# Strong commonality

*Strong* commonality: **intersection** of features observed in examples, **minus union** of features in counter-examples

Example #1:
He's going to be rich one day.

Example #2:
I'm going to think about it.



Counter-example #1:
He will be a rich man one day.

Counter-example #2:
It was when we talked about the meeting.

# Commonality – case study #2

Another case study: the imperative mood, e.g.

- Be that way.
- Dream a little dream of me.
- Think about the children.

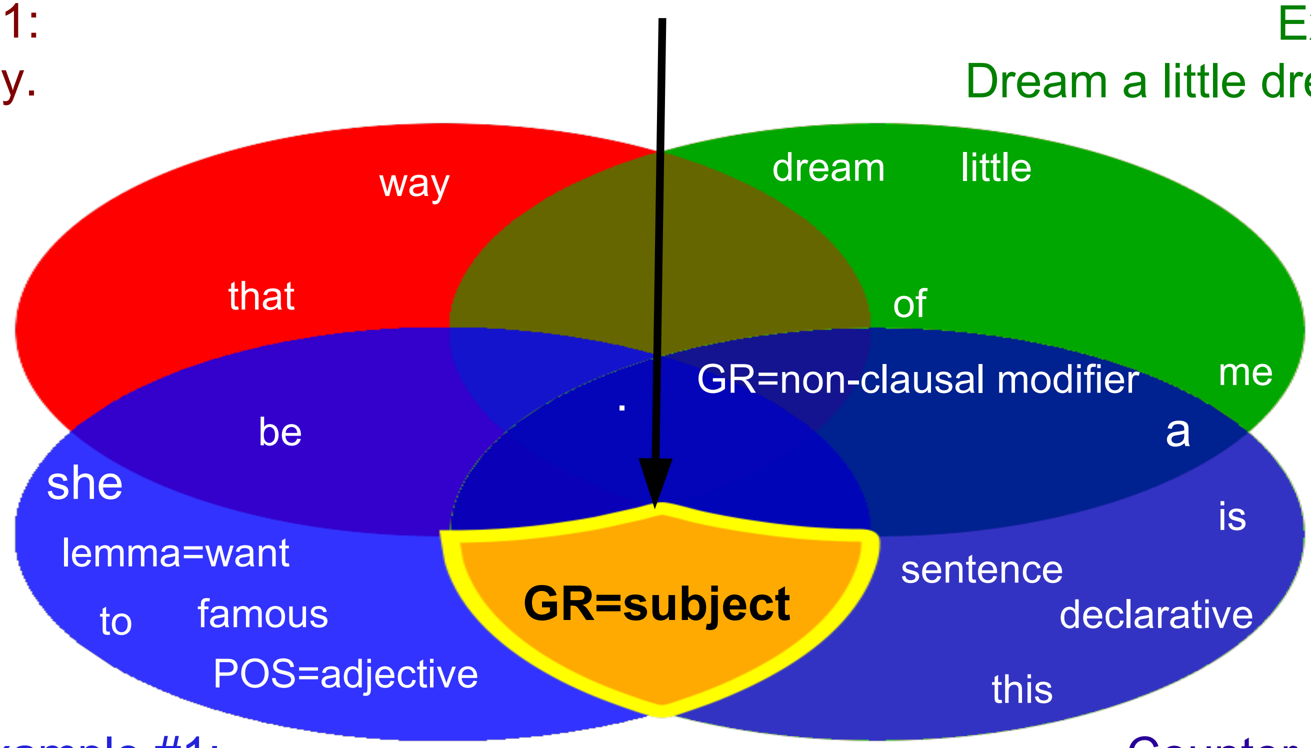
Can't spot imperative using approaches described so far...

# Absence-of commonality

*Absence-of* commonality: **intersection** of features observed in **counter-examples**, minus union of features in **examples**

Example #1:
Be that way.

Example #2:
Dream a little dream of me.



Counter-example #1:
She wanted to be famous.

Counter-example #2:
This is a declarative sentence.

# But it's not enough...

- This much works in a few simple cases, but:
  - False positive: “I'm going to the shops to buy some tea”
  - **Composite features** (e.g. a noun that is the subject of a verb) far more likely to capture commonality...

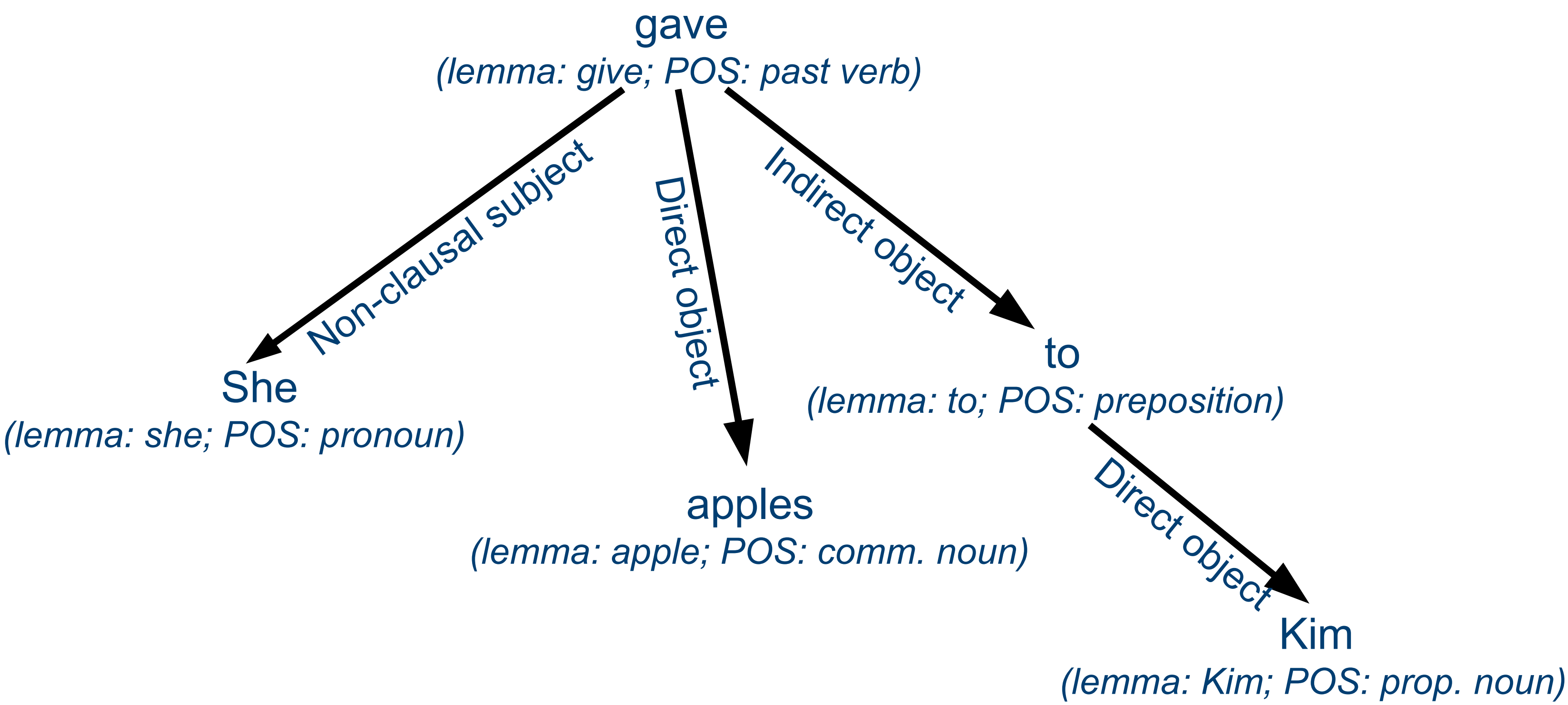
# Analysing a sentence again – dependency structures

GRs + words = *almost* a tree...

(directed graph, but sometimes cyclic)

*(Thank you, Google Image Search)*

# Analysing a sentence again – dependency structures



gave
*(lemma: give; POS: past verb)*

Non-clausal subject

Direct object

Indirect object

She
*(lemma: she; POS: pronoun)*

apples
*(lemma: apple; POS: comm. noun)*

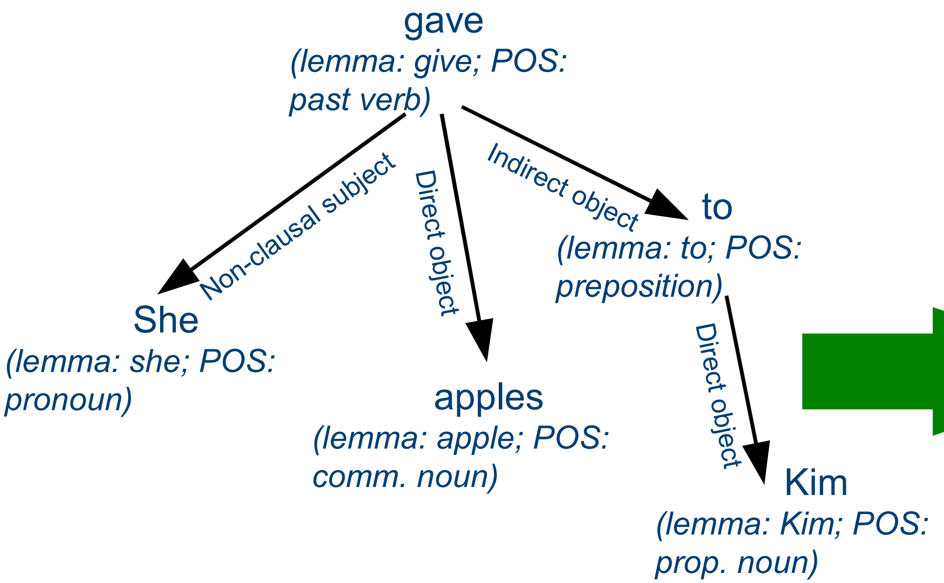
to
*(lemma: to; POS: preposition)*

Direct object

Kim
*(lemma: Kim; POS: prop. noun)*

# Break [almost-]tree up into smaller features: 'twigs'

# Break [almost-]tree up into smaller features: 'twigs'



gave
*(lemma: give; POS: past verb)*

Non-clausal subject

Direct object

Indirect object

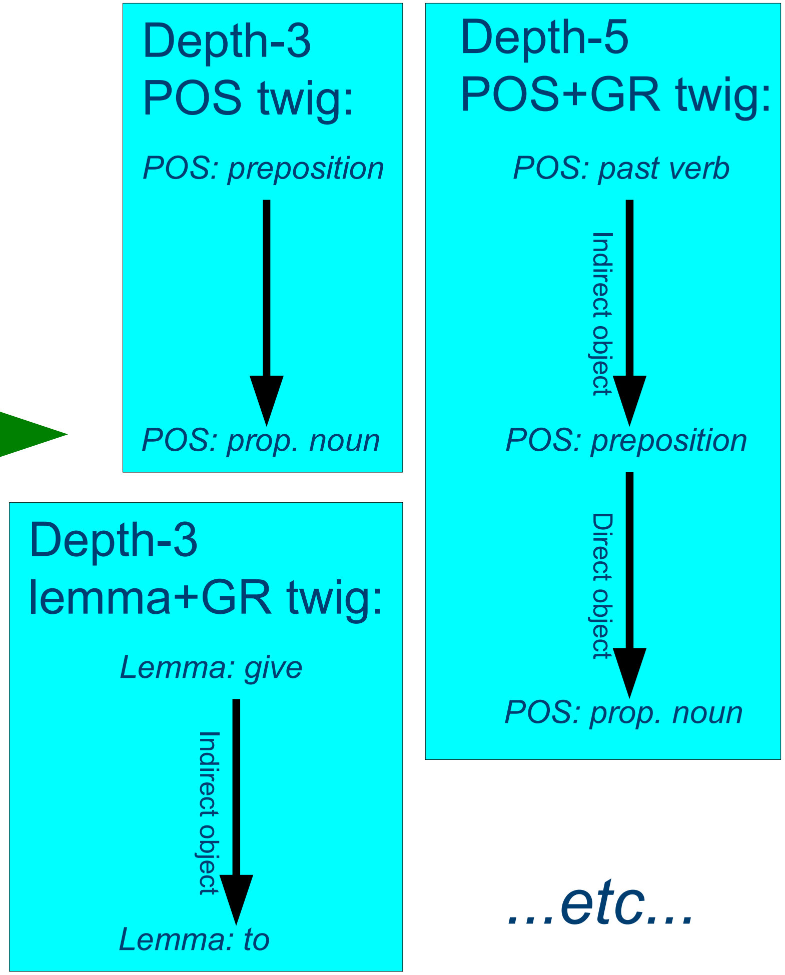
She
*(lemma: she; POS: pronoun)*

apples
*(lemma: apple; POS: comm. noun)*

to
*(lemma: to; POS: preposition)*

Direct object

Kim
*(lemma: Kim; POS: prop. noun)*



Depth-3 POS twig:

*POS: preposition*

*POS: prop. noun*

Depth-5 POS+GR twig:

*POS: past verb*

Indirect object

*POS: preposition*

Direct object

*POS: prop. noun*

Depth-3 lemma+GR twig:

*Lemma: give*

Indirect object

*Lemma: to*

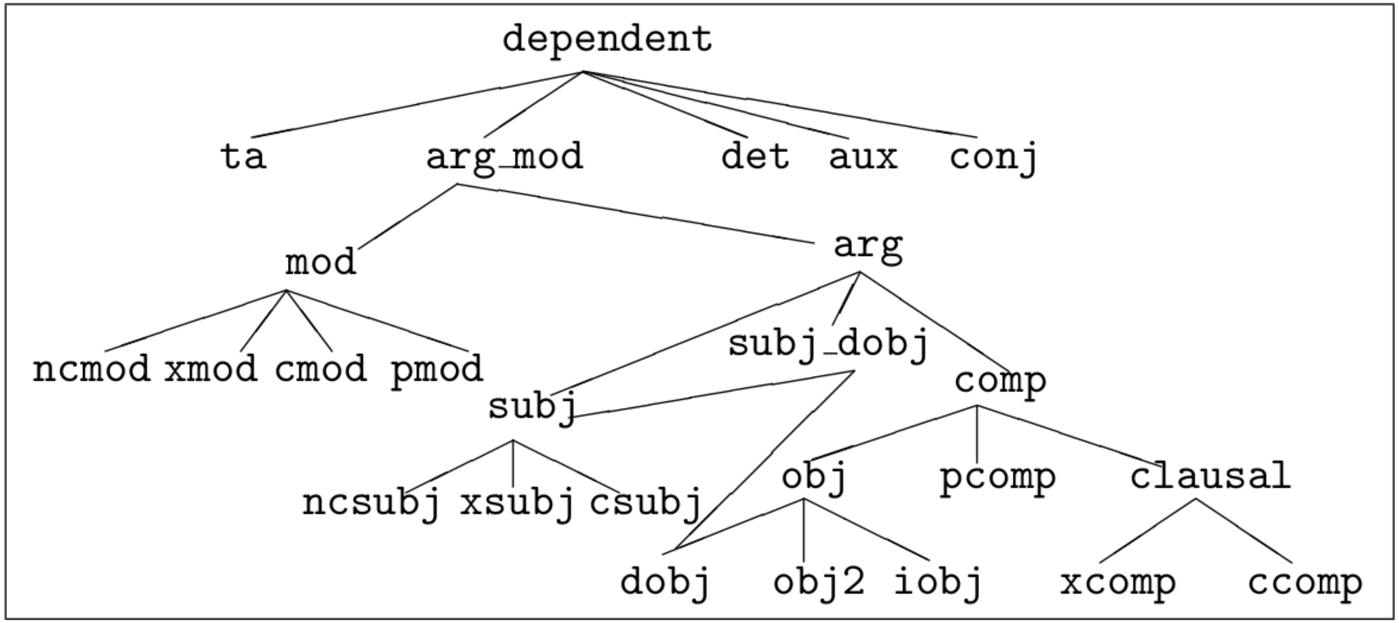
1. Break each example's graph into 'twigs'
2. Check which twigs are common to all examples

*...etc...*

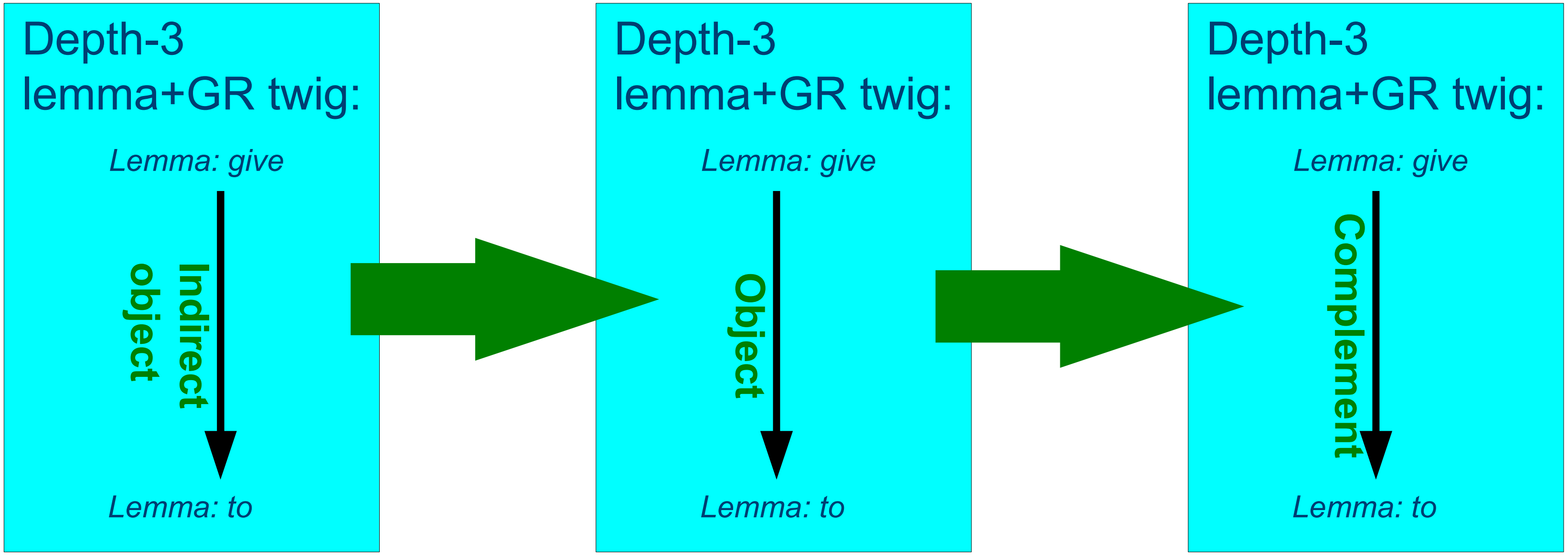
UNIVERSITY OF CAMBRIDGE

# Sometimes this is still not enough...

- What if our learning task is transitive verbs, but we're not interested in the direct/indirect object distinction?
  - Use GR type hierarchy to generate features



# Exploiting the GR type hierarchy



- Go up GR type hierarchy, generate new feature at each level
- Might only find commonality between example sentences at a higher level
- Same approach for POS tags, but requires custom hierarchy: standard Penn-Treebank set is flat

# Running *LearnGrammar!* on Android (Nexus One)

- It works! (Mostly. Formal evaluation outstanding)
- Corpus load is slow
  - ~45s for 1000 pre-parsed sentences after mucking with file format
  - Run in background thread while user enters sentences
- Example sentence parsing via webservice call
- Searching for similar sentences is currently ~85s
  - Heavy pruning of redundant features required
  - Planning on implementing lazy search to avoid long up-front pause