CENTRE FOR LANGUAGE TECHNOLOGY
MACQUARIE UNIVERSITY – SYDNEY

At Last! A Reason to Generate Language from Logic

Robert Dale
rdale@ics.mq.edu.au

Collaborative work with Dave Barker-Plummer, Stanford; Richard Cox, U Sussex; Mark Dras and Rolf Schwitter, Macquarie U

The Aims of This Talk

- To introduce a new problem in Natural Language Generation
- To sketch the approach we intend to take
- To provide some initial data analysis

Agenda

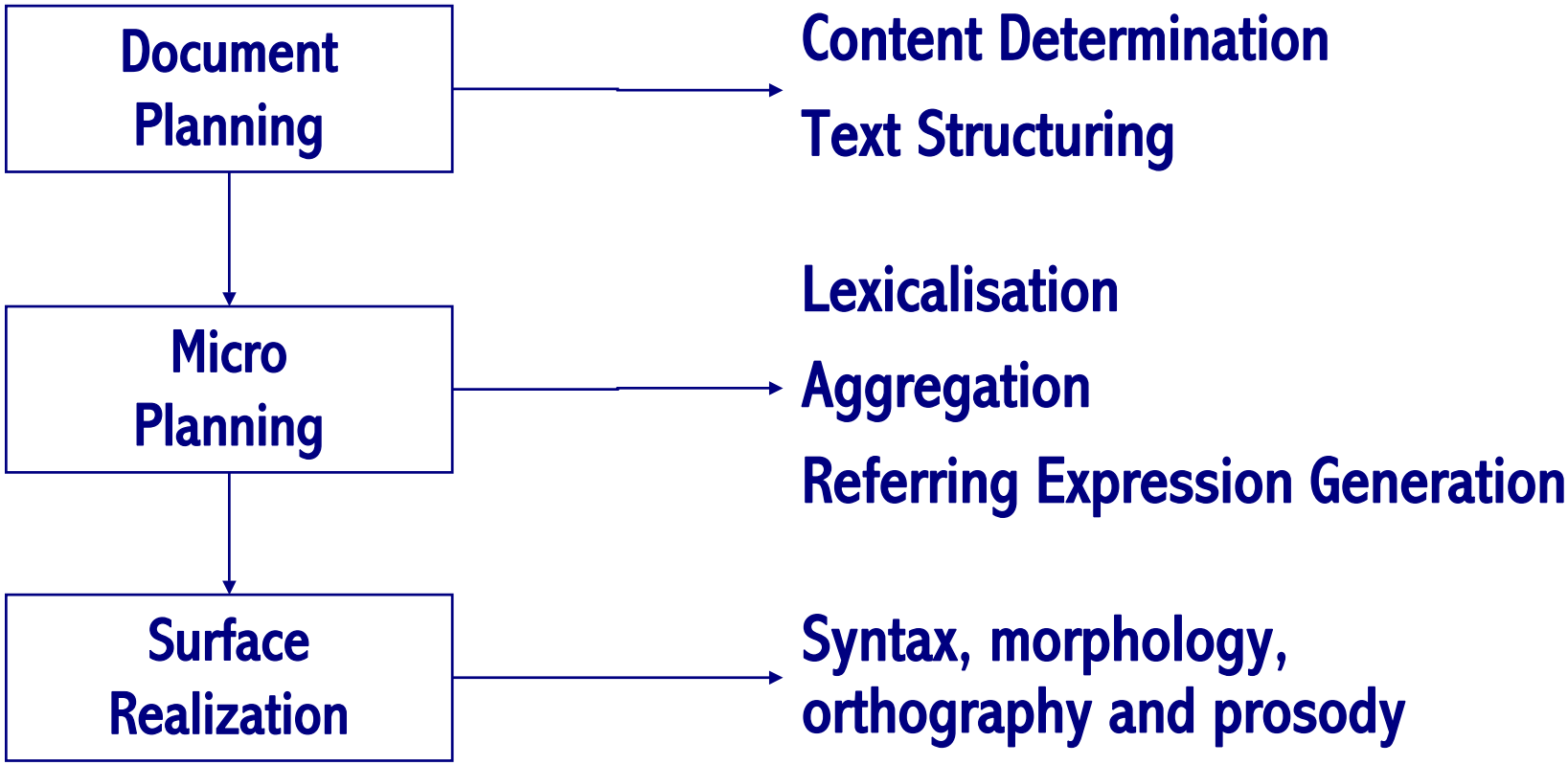
- Approaches to Generation, Past and Present
- The OpenProof Project
- Paraphrase Selection
- A Look at Some Real Data
- Next Steps

How Natural Language Generation Used To Be Done

The predominant approach until this decade:

- Requires a rich input knowledge representation
- Discourse generation starts with a communicative goal
- Makes subtle linguistic decisions about what to say and how to say it using a domain model, a discourse model and a user model

A Traditional NLG Architecture

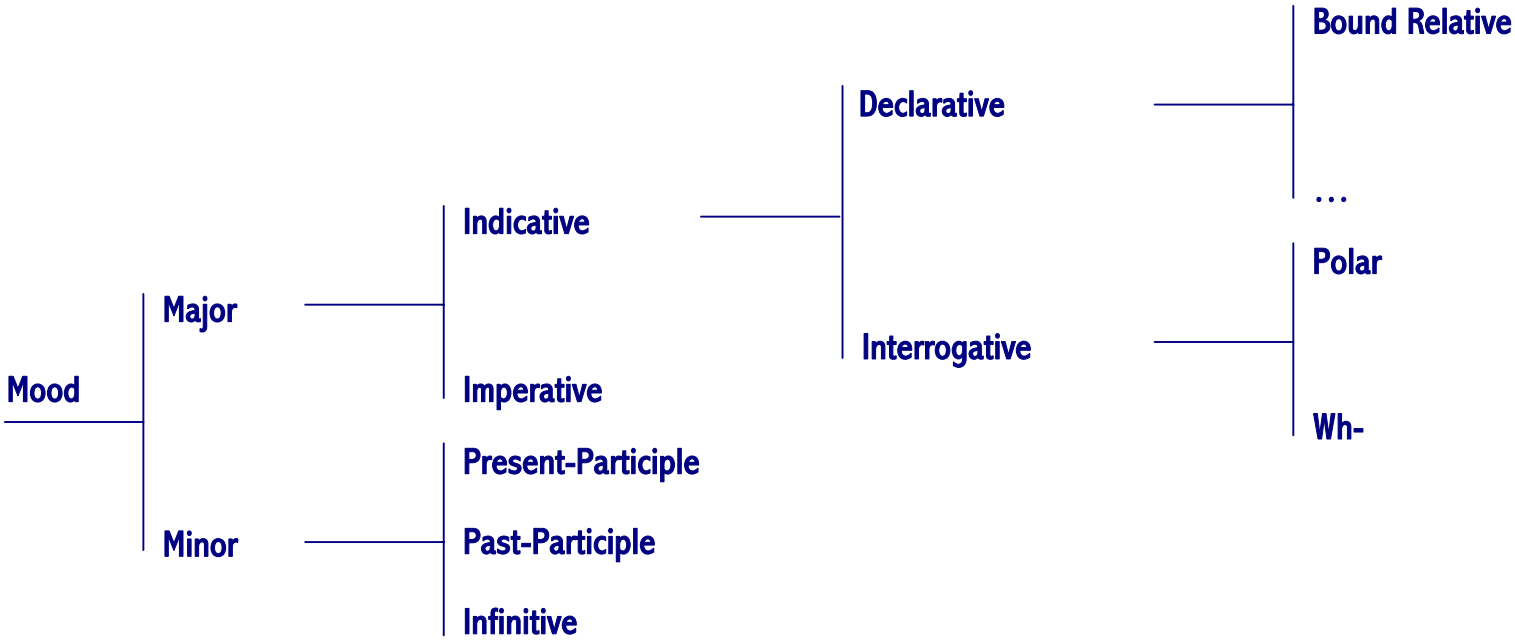


One Example:
An SPL input to KPML

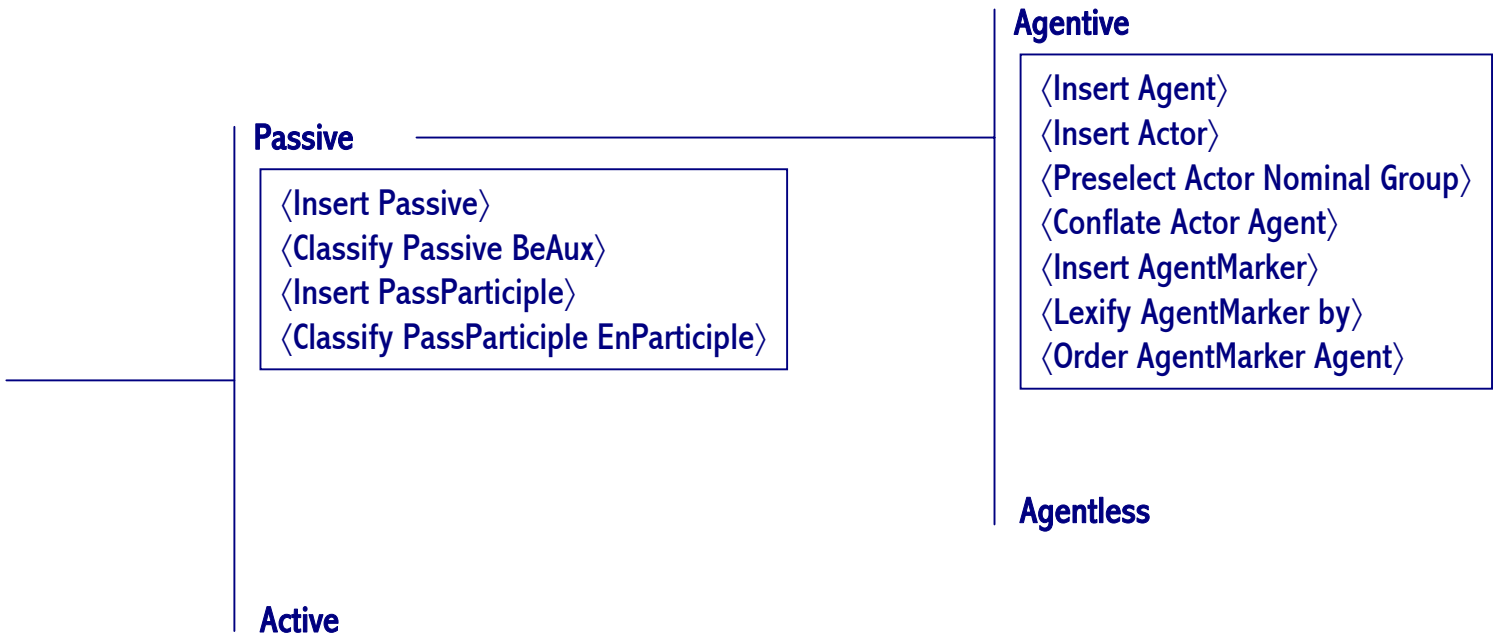
```
(l / greater-than-comparison
  :tense past
  :exceed-q (l  a) exceed
  :command-offer-q notcommandoffer
  :proposal-q notproposal
  :domain (m / one-or-two-d-time :lex month :determiner the)
  :standard (a / quality :lex average determiner zero)
  :range (c / sense-and-measure-quality :lex cool)
  :inclusive (r / one-or-two-d-time
    :lex day
    :number plural
    :property-ascription (r / quality :lex rain)
    :size-property-ascription
                        (av / scalable-quality :lex the-av-no-of)))

The month was cooler than average with the average number of rain days.
```

Decision Making in a Systemic Network



Realisation Statements



How Natural Language Generation Gets Done Today

- Input is either:
  - an underspecified knowledge representation
  - other texts
- Language models are used to choose most likely realisation

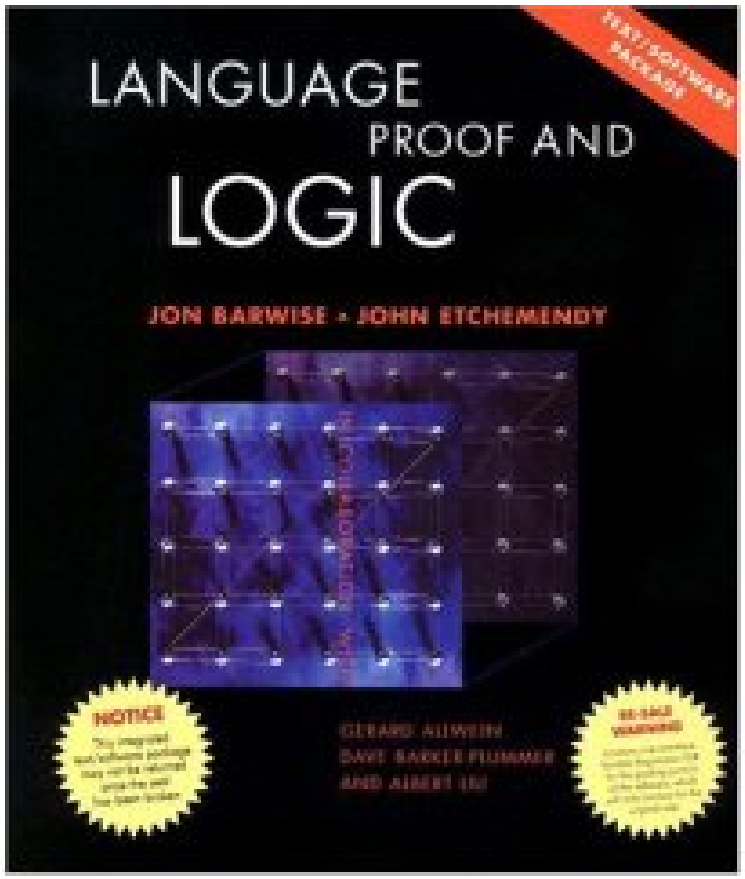
Problems

- For the earlier approaches:
  - The rich underlying representations just don't exist
- For the later approaches:
  - No insights into the really interesting questions about language use

Agenda

- Approaches to Generation, Past and Present
- The OpenProof Project
- Paraphrase Selection
- A Look at Some Real Data
- Next Steps

Language, Proof and Logic



A Translation Exercise

7.12 (Translation) Translate the following English sentences into FOL. Your translations will use all of the propositional connectives.

1. *If* $\boldsymbol{a}$ *is a tetrahedron then it is in front of* $\boldsymbol{d}$.
2. $\boldsymbol{a}$ *is to the left of or right of* $\boldsymbol{d}$ *only if it's a cube.*
3. $\boldsymbol{c}$ *is between either* $\boldsymbol{a}$ *and* $\boldsymbol{e}$ *or* $\boldsymbol{a}$ *and* $\boldsymbol{d}$.
4. $\boldsymbol{c}$ *is to the right of* $\boldsymbol{a}$, *provided it (i.e.,* $\boldsymbol{c}$*) is small.*

A Grade Grinder Report

EXERCISE-7.12.Sentences-7.12.error.1=*** Your first sentence, "FrontOf(a,d)→ Tet(a)", is not equivalent to any of the expected translations.

The Grade Grinder Dataset

The Grade Grinder

- can process solutions to 489 of the 748 exercises in the LPL book
- has been used by more than 38000 individual students over the last eight years, from around 100 institutions in around a dozen countries
- has assessed approximately 1.8 million individual submissions (each of which can contain zero or more exercises)

Hypothesis

- Perhaps we can provide better feedback by translating the student's errored solution back into natural language, so they can see their error

An Example

- English sentence:
  - John is either at the library or at home.
- Incorrect student translation (too weak):
  - $Lib(j) \vee Home(j)$
- Correct translation:
  - $Lib(j) \vee Home(j) \wedge \neg(Lib(j) \wedge Home(j))$
- A possible back-translation of the student's answer:
  - John is either at home or at the library or both.

What This Might Look Like

You were asked to translate:	John is either at the library or at home.
You translated this as:	$Lib(j) \vee Home(j)$
But what you said really means:	John is either at home or at the library or both.

Agenda

- Approaches to Generation, Past and Present
- The OpenProof Project
- Paraphrase Selection
- A Look at Some Real Data
- Next Steps

Generating Paraphrases

The Basic Idea:

- The same logical form can be rendered in many different ways in NL
- Some renderings may be easier for a student to understand
- Some renderings may make it easier for a student to see where they have gone wrong

The Aim:

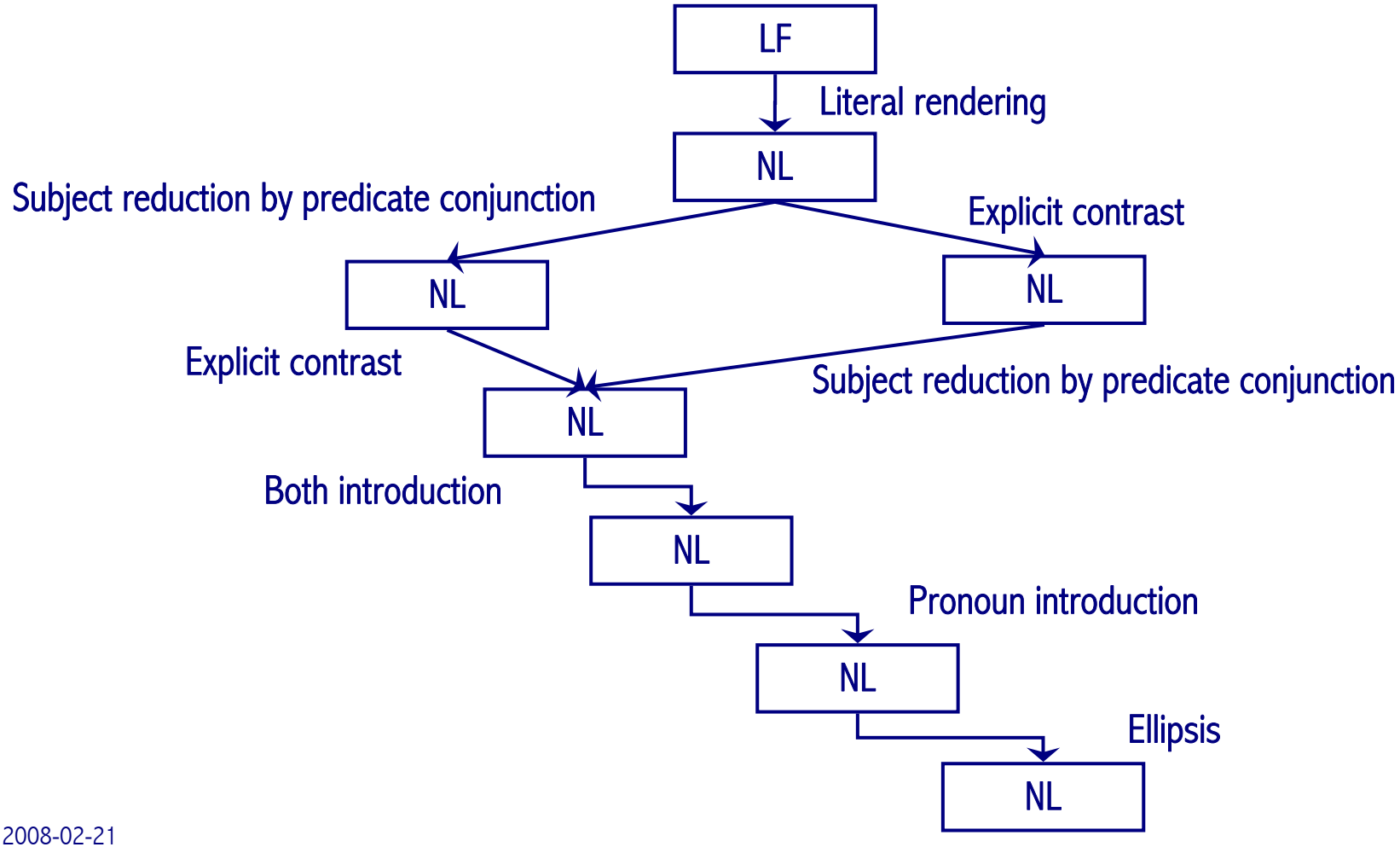
- to develop automatic natural language paraphrase capabilities that, given a student's incorrect answer, are able to select and formulate an appropriate natural language expression that makes clear the difference between this and the correct answer

Paraphrase 'Distance From Source'

$[Home(john) \vee Home(mary)] \wedge \neg[Home(john) \wedge Home(mary)]$

- Either John is home or Mary is home and it's not the case that John is home and Mary is home
- Either John or Mary is home and it's not the case that John and Mary are both home
- Either John or Mary is home but it's not the case that John and Mary are both home
- Either John or Mary is home but it's not the case that both of them are home
- Either John or Mary is home but not both

A Paraphrase Graph



Basic Ideas

- Paraphrase n is rewritten as Paraphrase m by a tree rewrite rule
- Rewrite rules have a cost, or cause a certain amount of damage (including information loss)
- Paraphrases have properties or effects: they emphasise certain things
- The further a paraphrase is from the literal rendering the harder it may be to see the relationship between logic and NL …
- … but literal renderings can be significantly more complex than the simplest NL rendering

Paraphrases #2

- $\forall x \forall y \forall z$ ((FatherOf(x,y) $\wedge$ FatherOf(y,z)) $\rightarrow$ Nicer(x,y)
- For all x, y and z, if x is the father of y and y is the father of z then x is nicer than y
- For all x, y and z, if x is z's paternal grandfrather and y is z's father, then x is nicer than y
- For all z, z's paternal grandfather is nicer than z's father
- It's the case for everyone that their paternal grandfather is nicer than their father

Paraphrases #3: De Morgan's Laws

- $\neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q$
  - It's not the case that both P and Q ⇔ Either not P or not Q
  - It's not the case that both John and Simon are telling the truth
  - Either John isn't telling the truth or Simon isn't telling the truth
- Add 'synonymy by negation':
  - Either John is lying or Simon is

Contextual Constraints on Paraphrase Choice

What we know or might be able to infer:

- The specific mistake that has been made
- The extent to which the student is comfortable with other parts of the translation
- What concepts they are already comfortable with
- What mistakes they have made before

So:

- Learn the mapping from user model and task model to preferred paraphrase

Agenda

- Approaches to Generation Past and Present
- The OpenProof Project
- An Approach to Paraphrase Selection
- Some Data Analysis
- Next Steps

Data Selection for Initial Exploration

- We computed the number of GG submissions per LPL exercise and rank ordered them; Exercise 7.12 from Chapter 7 (which introduces conditionals) was selected
- 74,000 submitted solutions, of which 42,416 were erroneous (57%), containing 148,681 incorrect translation solutions
- The solutions were submitted by 11,925 students representing an average of 12.47 erroneous sentences per student

Exercise 7.12: Sentences 1-10

1. If a is a tetrahedron then it is in front of d.
2. a is to the left of or right of d only if it's a cube.
3. c is between either a and e or a and d.
4. c is to the right of a, provided it (i.e., c) is small.
5. c is to the right of d only if b is to the right of c and left of e.
6. if e is a tetrahedron, then it's to the right of b if and only if it is also in front of b.
7. If b is a dodecahedron, then if it isn't in front of d then it isn't in back of d either.
8. c is in back of a but in front of e.
9. e is in front of d unless it (i.e., e) is a large tetrahedron.
10. At least one of a, c, and e is a cube.

Exercise 7.12: Sentences 11-20

11. a is a tetrahedron only if it is in front of b.
12. b is larger than both a and e.
13. a and e are both larger than c, but neither is large.
14. d is the same shape as b only if they are the same size.
15. a is large if and only if it's a cube.
16. b is a cube unless c is a tetrahedron.
17. If e isn't a cube, either b or d is large.
18. b or d is a cube if either a or c is a tetrahedron.
19. a is large just in case d is small.
20. a is large just in case e is.

An Error Taxonomy

45 distinct error types organised under the following categories:

- Structural Errors
- Connective Errors
- Atomic Errors
  - Predicate Errors
  - Argument Errors

Examples of Errors

#	Reference solution	Errored solution	Type	Subtype
1	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{FrontOf}(a,d) \rightarrow \mathsf{Tet}(a)$	1	Antecedent–Consequent Reversal
2	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{FrontOf}(a,b) \rightarrow \mathsf{Tet}(a)$	1 and 3ii	Antecedent–Consequent Reversal Incorrect Constant
3	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{Tet}(a) \vee \mathsf{FrontOf}(a,d)$	2	Disjunction for Conditional
4	$\neg\mathsf{Cube}(e) \rightarrow (\mathsf{Large}(b) \vee \mathsf{Large}(d))$	$\neg\mathsf{Cube}(e) \rightarrow \mathsf{Large}(b) \vee \mathsf{Large}(d)$	1	Missing Parens
5	$\mathsf{Large}(e) \rightarrow \mathsf{Large}(a)$	$e \rightarrow \mathsf{Large}(a)$	2	Elided Predicate
6	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{Tet}(a) \rightarrow \mathsf{InFrontOf}(a,d)$	3i	Incorrect Predicate
7	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,b)$	3ii	Incorrect Constant
8	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(a,d)$	$\mathsf{Tet}(a) \rightarrow \mathsf{FrontOf}(d)$	3ii	Arity Error

Error Frequencies

Error Type	Count	%age of All
Antecedent–Consequent Reversal	25084	25.86%
Biconditional for Conditional	17518	18.06%
Conditional for Biconditional	11362	11.71%
Negation Error	8954	9.23%
Incorrect Scope	5422	5.59%
Failure to Scope	4701	4.85%
Argument Error	4474	4.61%
Conjunction for Conditional	3187	3.29%
Conditional for Conjunction	2091	2.16%
Biconditional for Conjunction	1514	1.56%

BiCondForCond Errors

Frequency	Percentage	Surface Form
13214	75.43%	S only if S.
1777	10.14%	S unless S.
1146	6.54%	S provided S.
725	4.14%	S if S.
367	2.09%	If S then if S then S.
289	1.65%	If S then S.

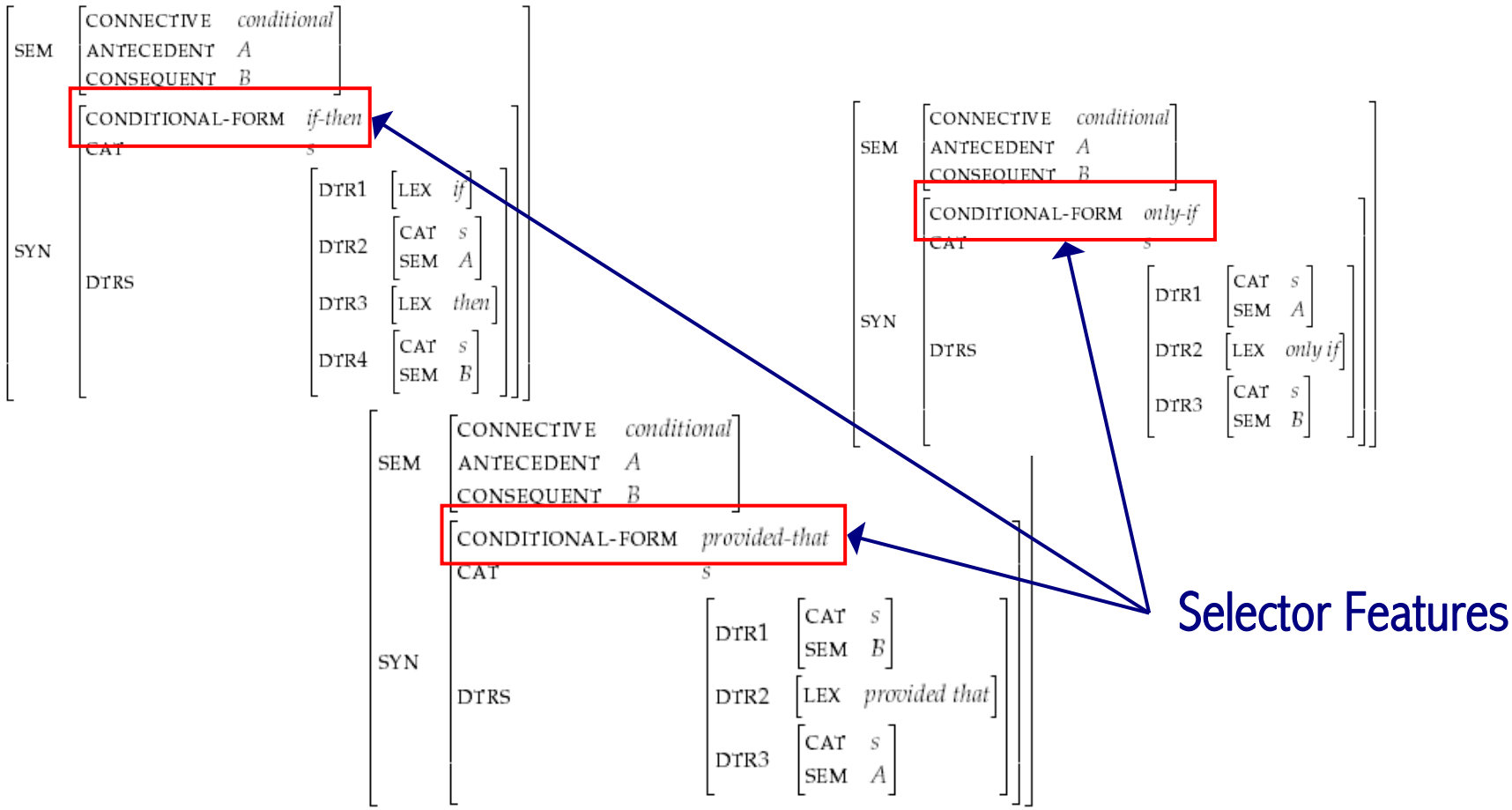
Agenda

- Approaches to Generation, Past and Present
- The OpenProof Project
- Paraphrase Selection
- A Look at Some Real Data
- Next Steps

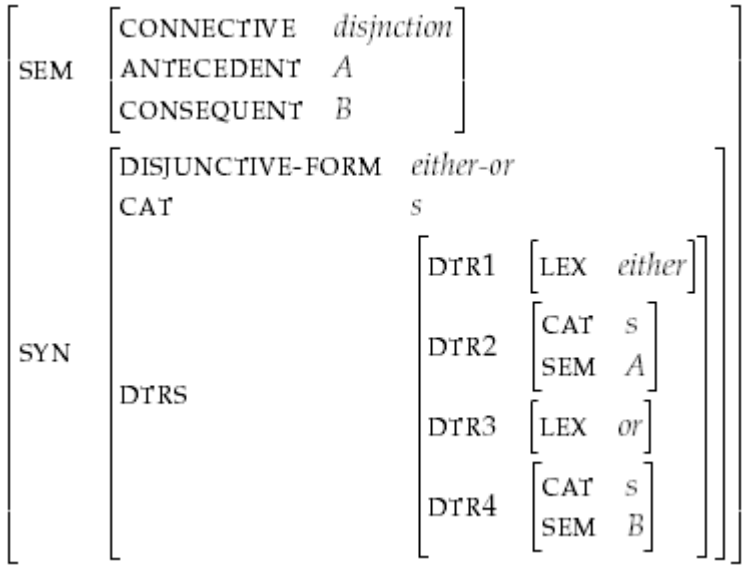
Logic to NL Correspondences

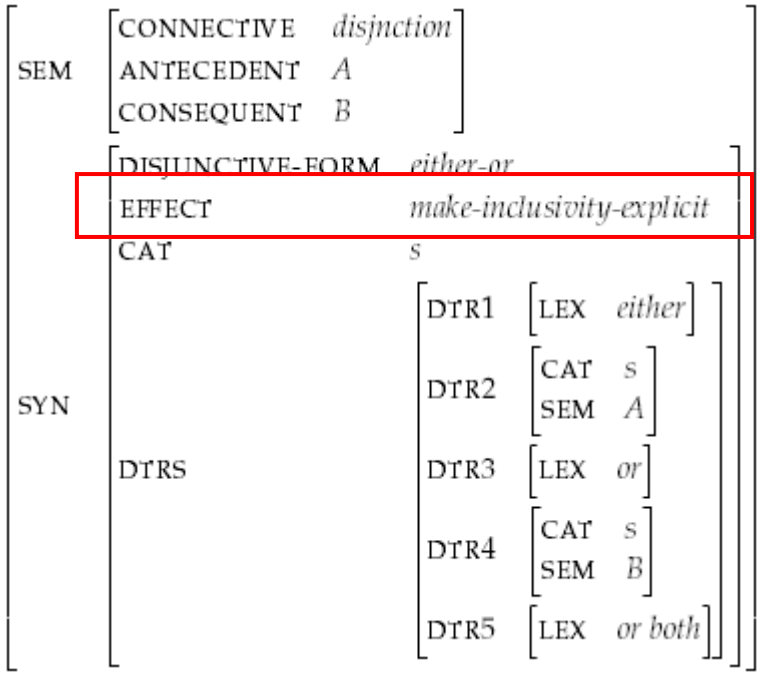
$$\begin{bmatrix} \text{SEM} & \begin{bmatrix} \text{CONNECTIVE} & \textit{conditional} \\ \text{ANTECEDENT} & A \\ \text{CONSEQUENT} & B \end{bmatrix} \\ \text{SYN} & \begin{bmatrix} \text{CONDITIONAL-FORM} & \textit{if-then} \\ \text{CAT} & s \\ \text{DTRS} & \begin{bmatrix} \text{DTR1} & \begin{bmatrix} \text{LEX} & \textit{if} \end{bmatrix} \\ \text{DTR2} & \begin{bmatrix} \text{CAT} & s \\ \text{SEM} & A \end{bmatrix} \\ \text{DTR3} & \begin{bmatrix} \text{LEX} & \textit{then} \end{bmatrix} \\ \text{DTR4} & \begin{bmatrix} \text{CAT} & s \\ \text{SEM} & B \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Realisation Classes: Different Realisations of the Conditional



Realisation Classes: Surface Form Effects





Generation Strategy

- Malrules detect the types of errors found in the student's solution
- Each malrule results in directives for the generator to select structures that have particular features
- In complex cases there may be conflicting requirements
  - The generator should try to select the combination of features most likely to result in understanding
  - Best choice determined by weightings derived from the user and task model

Next Steps

- Further development of the error taxonomy and malrules
- Characterisation of a range of paraphrase rules to deal with the common cases
- Implementation of a prototype generator

Conclusions

- Traditional NLG requires:
  - a rich semantic input representation to motivate linguistic distinctions
  - widely varying contexts of use to motivate variation in output
- OpenProof + an immense student base provides both
- Other possibilities for the same approach:
  - Tailored advice in language learning
  - Customised web pages based on browsing history