

Outline for today

- Grammars
- First-order logic
- Learning a theory and new concepts in first-order logic.

Finite-state grammar

- The “minimal linguistic theory”.

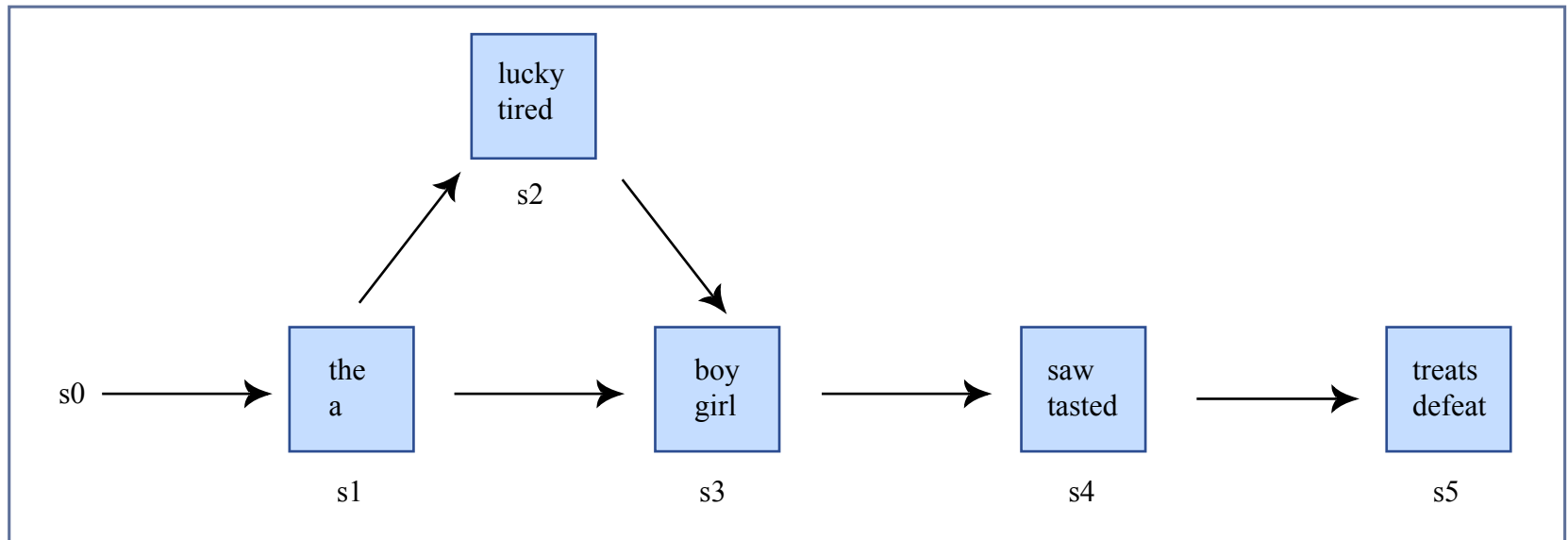


Figure by MIT OCW.

E.g., “The lucky boy tasted defeat.”

Generating infinite strings

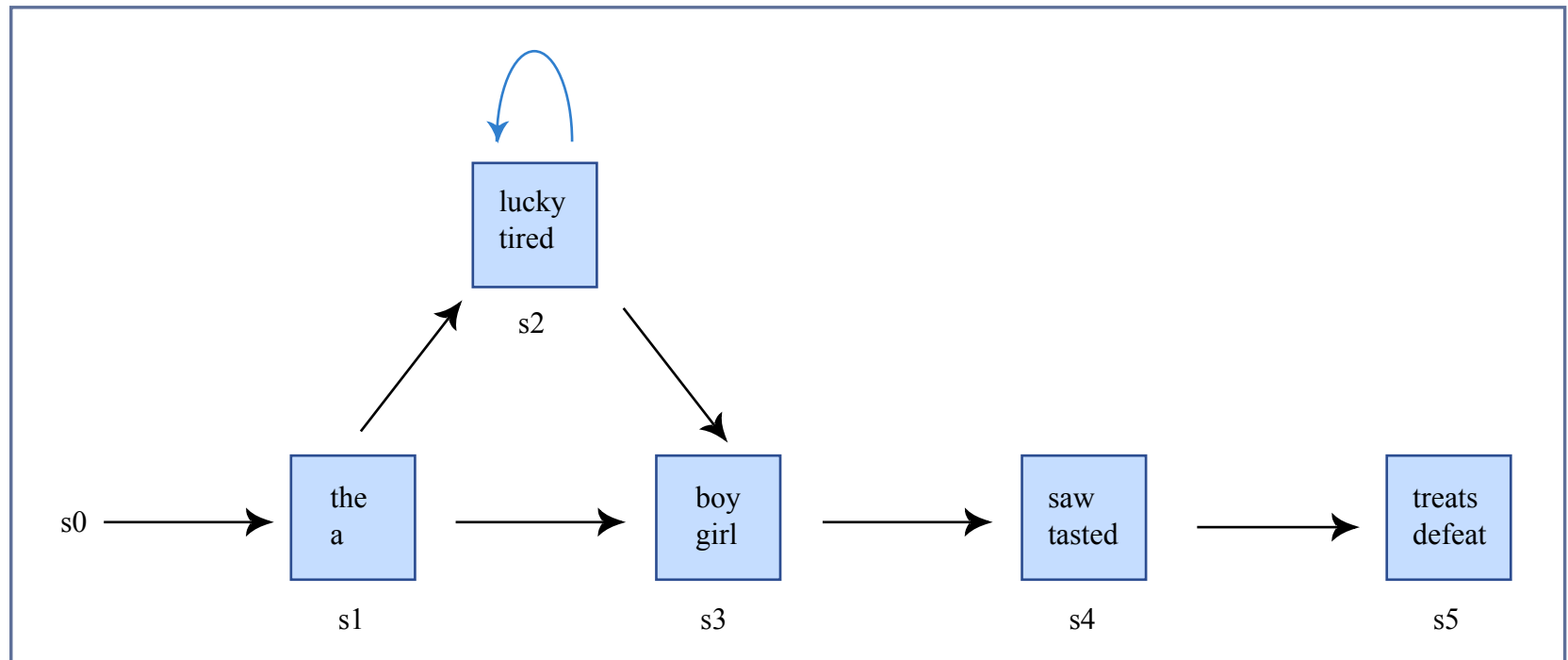


Figure by MIT OCW.

E.g., “The lucky tired tired ... boy tasted defeat.”

Parses novel sequences

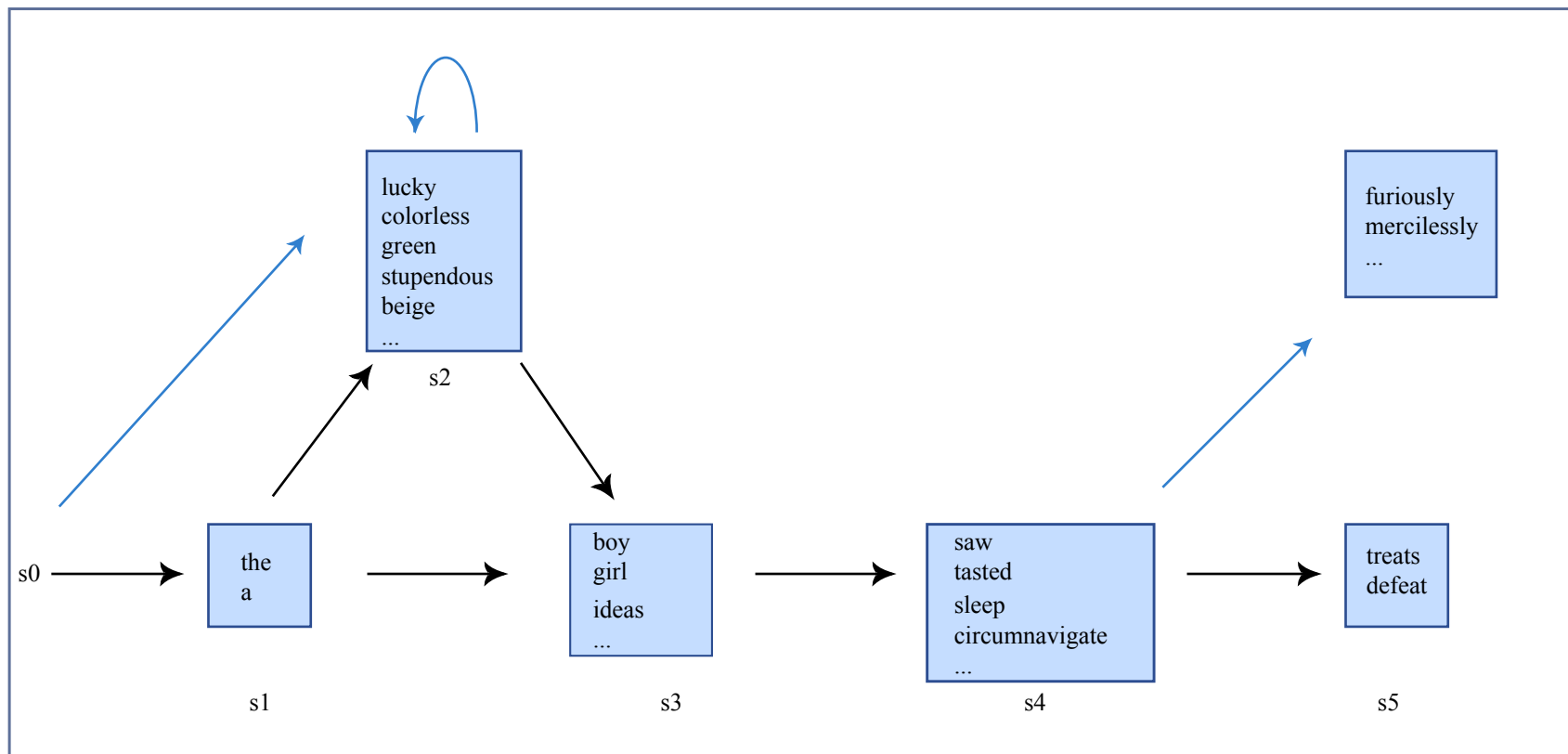


Figure by MIT OCW.

E.g., “Colorless green ideas sleep furiously”

What's really wrong

- The problem with the “statistical” models isn't that they are statistical.
- Nature of representation:
 - N-grams: Perceptual/motor/superficial/concrete
 - Utterance is a *sequence of words*.
 - Chomsky: Cognitive/conceptual/deep/abstract
 - Finite-state grammar: utterance is a *sequence of states*.
 - Phrase-structure grammar: utterance is a *hierarchical structure of phrases*.

Counterexamples to sequential models

- Center embedding
 - The man dies.
 - The man that the racoons bite dies.
 - The man that the racoons that the dog chases bite dies.
 - The man that the racoons that the dog that the horses kick chases bite dies.

Generating embedded clauses

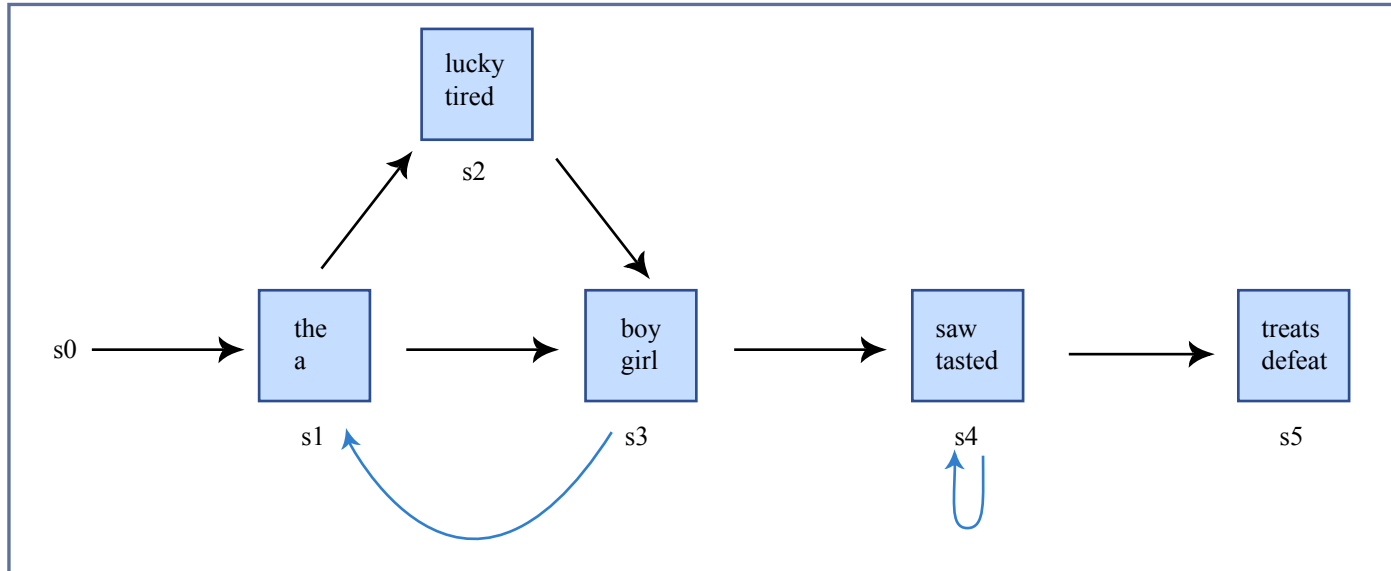


Figure by MIT OCW.

E.g., “A lucky boy the tired girl saw tasted defeat.”

No way to ensure subject-predicate agreement.

Generating embedded clauses

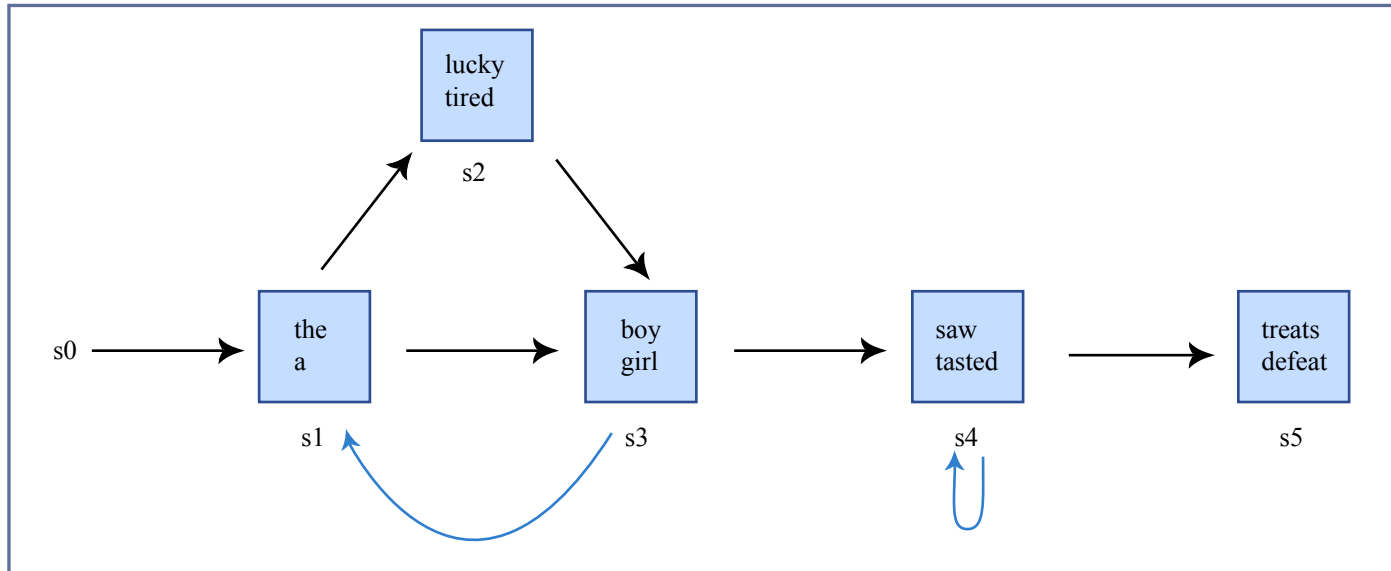


Figure by MIT OCW.

E.g., “A lucky boy the tired girl saw tasted defeat.”

“A lucky boy the tired girl saw saw saw saw ... saw treats.”

Counterexamples to sequential models

- Tail-embedding
 - The horse that kicked the dog that chased the racoons that bit the man is alive.
 - The horses that kicked the dog that chased the racoons that bit the man are alive.
- Fundamental problem: Dependencies may be arbitrarily long-range in the sequence of words. Only local in the *phrase structure*.

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

“Context-free” or “Phrase-structure” grammar:

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Image removed due to copyright considerations. Please see:

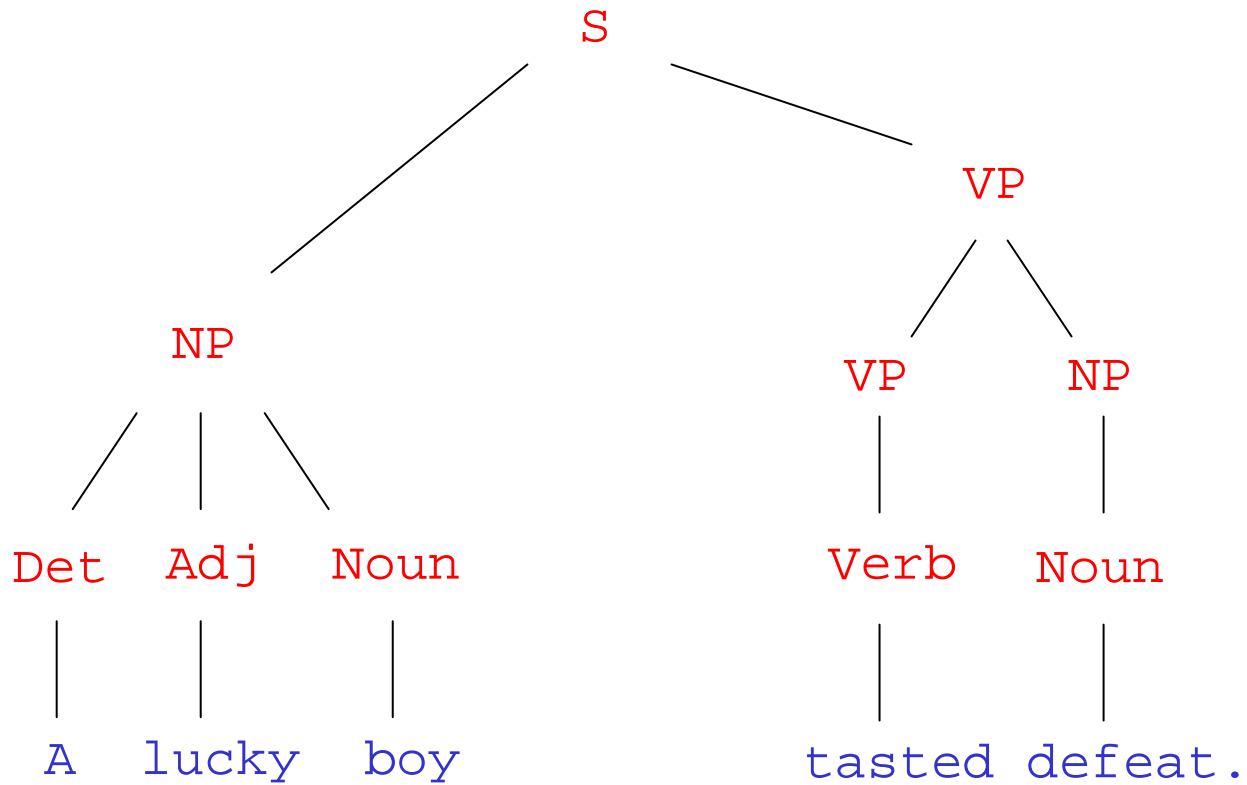
<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Context-free grammars for embedded clauses

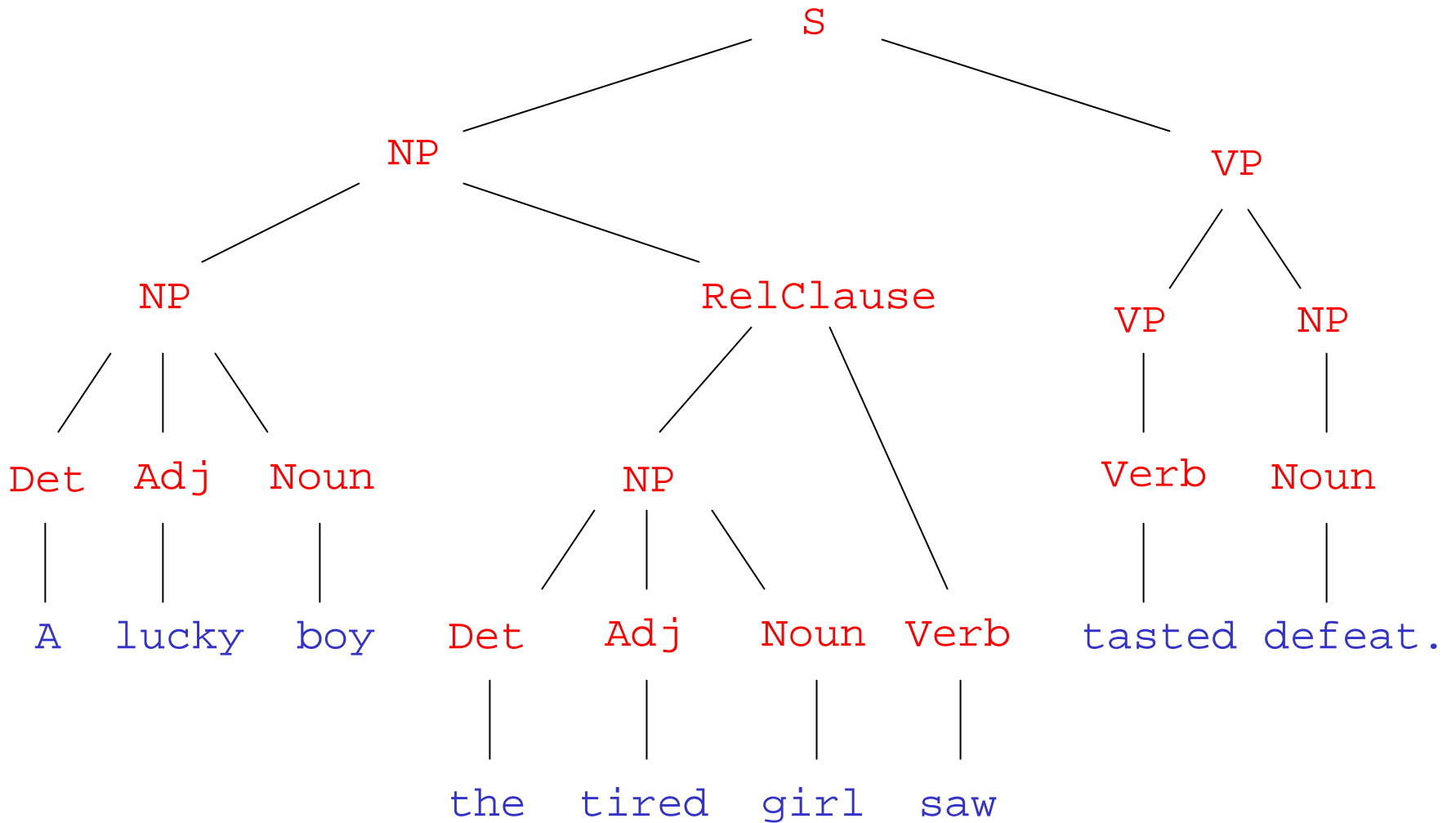
- For center embedding:

$$S \rightarrow NP VP$$
$$NP \rightarrow Det [Adj] Noun [RelClause]$$
$$RelClause \rightarrow [Rel] NP V$$
$$VP \rightarrow VP NP$$
$$VP \rightarrow Verb$$

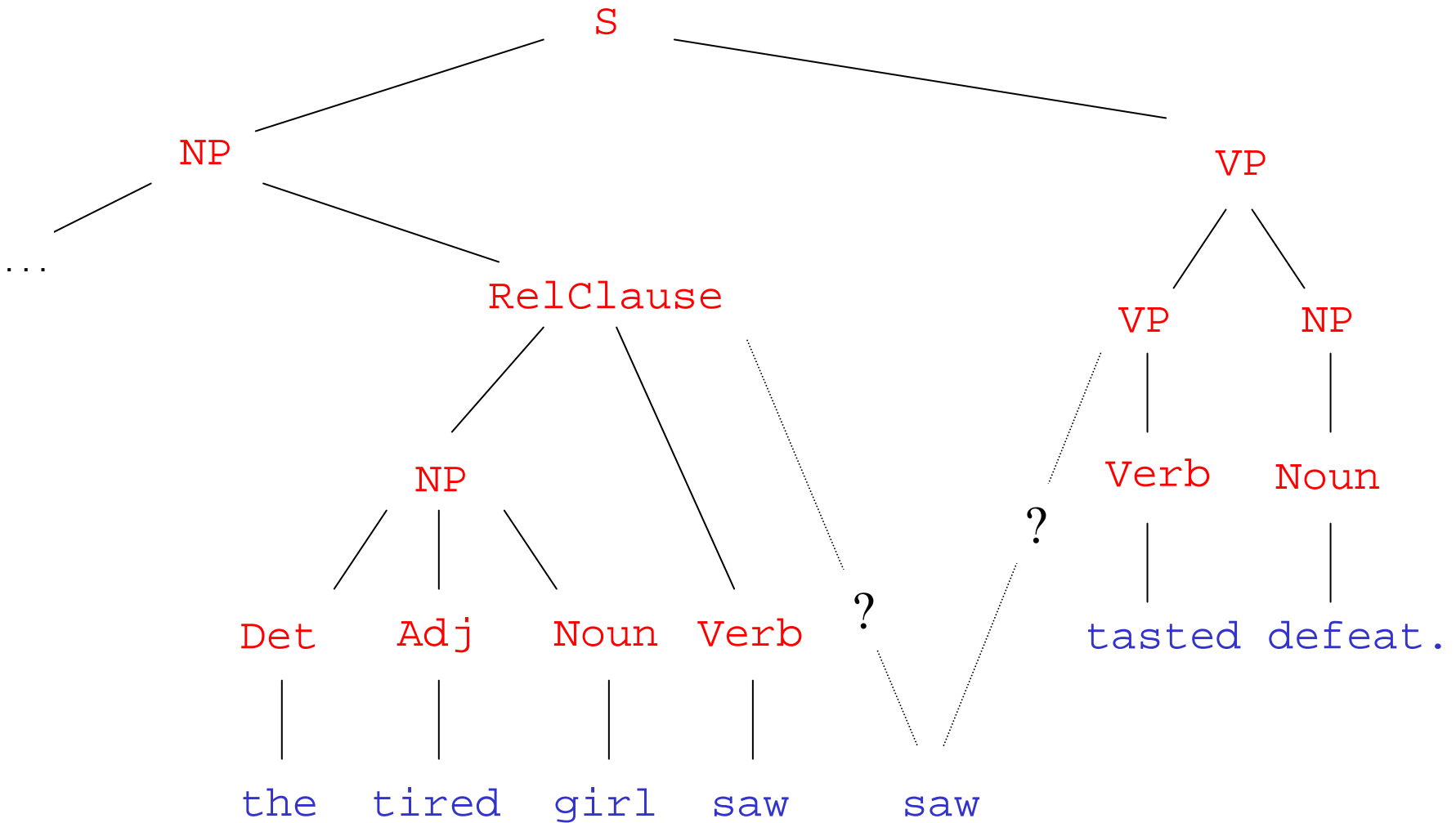
Phrase structure



Phrase structure



Phrase structure



What is special about this “structured” representation?

- Generative:
 - “Infinite use of finite means”
- Recursive.
 - E.g., for any sentence S, can say, “I think that S”, “Mary desires that S”, “Napoleon’s Italian stepmother naively hoped that S”,
- Generates structures of greatly varying complexity.
- Abstract.
 - Principles formulated over abstract concepts (e.g., NP, VP), whose meaning is defined only within the system.
- Multiple levels of abstraction....

Class of grammars



Grammar



Phrase structure



Utterance

Context-free grammar:

- Terminals
- Nonterminals
- Rewrite rules: One terminal -> one or more terminals or nonterminals

$$S \rightarrow NP VP$$
$$NP \rightarrow Det [Adj] Noun [RelClause]$$
$$RelClause \rightarrow [Rel] NP V$$
$$VP \rightarrow VP NP$$
$$VP \rightarrow Verb$$

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Class of grammars



Grammar



Phrase structure



Utterance

parsing

Context-free grammar:

- Terminals
- Nonterminals
- Rewrite rules: One terminal \rightarrow one or more terminals or nonterminals

$S \rightarrow NP VP$

$NP \rightarrow Det [Adj] Noun [RelClause]$

$RelClause \rightarrow [Rel] NP V$

$VP \rightarrow VP NP$

$VP \rightarrow Verb$

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Class of grammars



Grammar



Phrase structure



Utterance

grammar
acquisition



Context-free grammar:

- Terminals
- Nonterminals
- Rewrite rules: One terminal -> one or more terminals or nonterminals

$S \rightarrow NP VP$

$NP \rightarrow Det [Adj] Noun [RelClause]$

$RelClause \rightarrow [Rel] NP V$

$VP \rightarrow VP NP$

$VP \rightarrow Verb$

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

SEQUITUR

Induction of context-free grammars based on two principles of compression:

1. No digram (sequence of two symbols) should appear twice.
2. Every rule should be used at least twice.

Image removed due to copyright considerations.

Image removed due to copyright considerations.

What about statistics?

- Probabilities can be added to any of these structured grammars.
 - Probabilistic finite-state automaton:
 $P(\text{state } i | \text{state } j)$, $P(\text{word } i | \text{state } j)$ instead of
 $P(\text{word } i | \text{word } j)$
 - Probabilistic context-free grammar:
 $P(\text{rewrite rule } i | \text{non-terminal } j)$

What about statistics?

- Probabilities can be added to any of these structured grammars.
 - Probabilistic finite-state automaton
 - Probabilistic context-free grammar
- Advantages:
 - Resolve ambiguities in parsing and translation. (e.g., Keller)
 - Integrate with speech recognition and visual perception.
 - Grammar induction as statistical learning.

Class of grammars

Context-free grammar:

- Terminals
- Nonterminals
- Rewrite rules: One terminal -> one or more terminals or nonterminals

↓
Grammar

$S \rightarrow NP VP$
 $NP \rightarrow Det [Adj] Noun [RelClause]$
 $RelClause \rightarrow [Rel] NP V$
 $VP \rightarrow VP NP$
 $VP \rightarrow Verb$

↓
Phrase structure

↓
Utterance



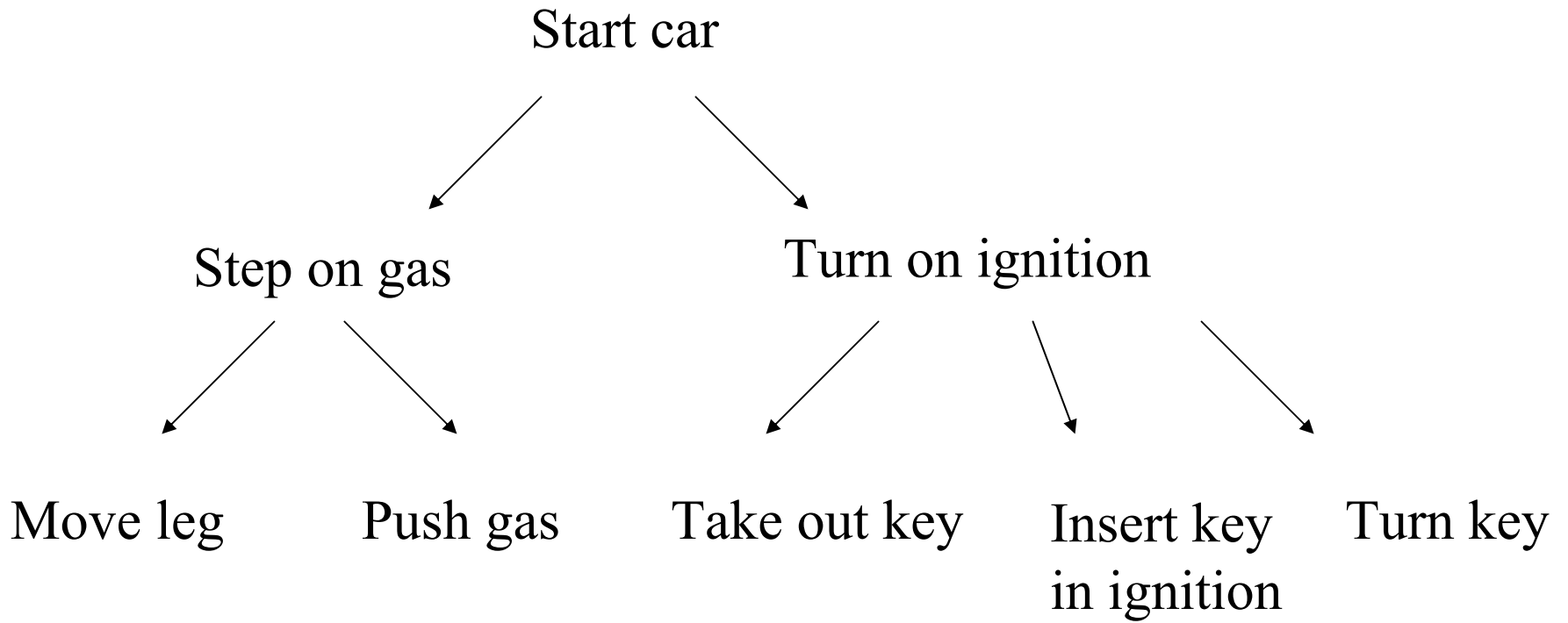
Image removed due to copyright considerations. Please see:
<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

Lessons

- Structure matters.
 - Can't get to phrase structure with purely sequential representations, no matter how powerful your statistics.
- Structure and statistics are complementary.
 - Statistics is about inference, structure is about representation.
 - Generative linguistics: negative example.
 - Modern computational linguistics: + example.
- What's special about a structured representation?
 - Generative, recursive, abstract

Other applications for phrase structures and grammars?

- Driving



Other applications for phrase structures and grammars?

- Driving
- Cooking

Image removed due to copyright considerations. Please see:
Forde, E. M. E., and G. W. Humphreys. "Category Specific Recognition
Impairments: A Review of Important Case Studies and Influential Theories."
Aphasiology 13, no. 3 (1999): 169-193.

Image removed due to copyright considerations. Please see:
Cooper, R. P., and T. Shallice. "Contention Scheduling and the Control of Routine Activities." *Cognitive Neuropsychology* 17 (2000): 297-338.

Other applications for phrase structures and grammars?

- Driving
- Cooking
- Music and dance
 - Schenker
 - Bernstein
 - Lerdahl and Jackendoff

Other applications for phrase structures and grammars?

- Driving
- Cooking
- Music and dance
 - Schenker
 - Bernstein
 - Lerdahl and Jackendoff
- Is any behavior *not* hierarchically organized? Is any behavior merely finite state?

The architecture of human behavior (1950's-1970's)

- Chomsky, *Syntactic Structures*
- Newell and Simon: The General Problem Solver
- Miller, Galanter, and Pribram, *Plans and the structure of behavior.*
- Production systems: Anderson, Newell
 - If GOAL and CONDITIONS then ACTION

Outline for today

- Grammars
- First-order logic
- Learning a theory and new concepts in first-order logic.

Why not just spaces or features to represent concepts?

- Euclidean space is fine for representing basic similarity relations between concepts, or values of concepts on a small number of fixed dimensions.
- But what about more abstract “concepts about concepts”?
 - Point x is *on the line between* y and z .
 - The relation between a and b is *a good analogy* for the relation between c and d .
 - The point x is the *nearest neighbor* of y .

First-order logic

- A language for talking about things, their properties and their relations.
 - Specific facts about particular objects
 - General laws about kinds of objects and their relations.
- In other words, a language for expressing theories of the world.

First-order logic

- Basic ingredients:

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

The grammar of first-order logic

Image removed due to copyright considerations. Please see:

<http://www.cs.umd.edu/class/fall2003/cmsc421-0101/chapter22-6pp.pdf>

First-order logic in cognition

- May be skeptical, because people have a hard time reasoning logically.
 - If A then B, A, therefore B. OK.
 - If A then B, not-B, therefore not A. Huh?
- What about probabilities and graded inferences?
- The importance of logic is in *expressive representation*, not necessarily *deductive inference*.

First-order logic in cognition

- Natural language semantics
 - “Somebody’s gonna hurt someone before the night is through.”
 - “Everybody wants to touch somebody if it takes all night.” This is subtle....

$$\exists y \forall x (\text{wants}(x, \text{touch}(x, y)))$$

$$\forall x \exists y (\text{wants}(x, \text{touch}(x, y)))$$

$$\forall x (\text{wants}(x, \exists y \text{touch}(x, y)))$$