

## NL IAS Marches On

Student: Lauren M. Stuart | Advisors: V. Raskin & J. M. Taylor

Natural language processing (NLP) has many applications for information assurance and security; among them are:

- Better memorization of random data for passwords [1]
- Watermarking NL texts [1][2]
- Automated translation for secrecy [1] or processing large amounts of NL documents
- Selective information hiding based on security classification [1]
- Identification of insider threat and social engineering [3]

[1] Mikhail J. Atallah, Craig J. McDonough, Victor Raskin, and Sergei Nirenburg. 2001. Natural language processing for information assurance and security: an overview and implementations. In *Proceedings of the 2000 workshop on New security paradigms (NSPW '00)*. ACM, New York, NY, USA, 51-65. DOI=10.1145/366173.366190 <http://doi.acm.org/10.1145/366173.366190>

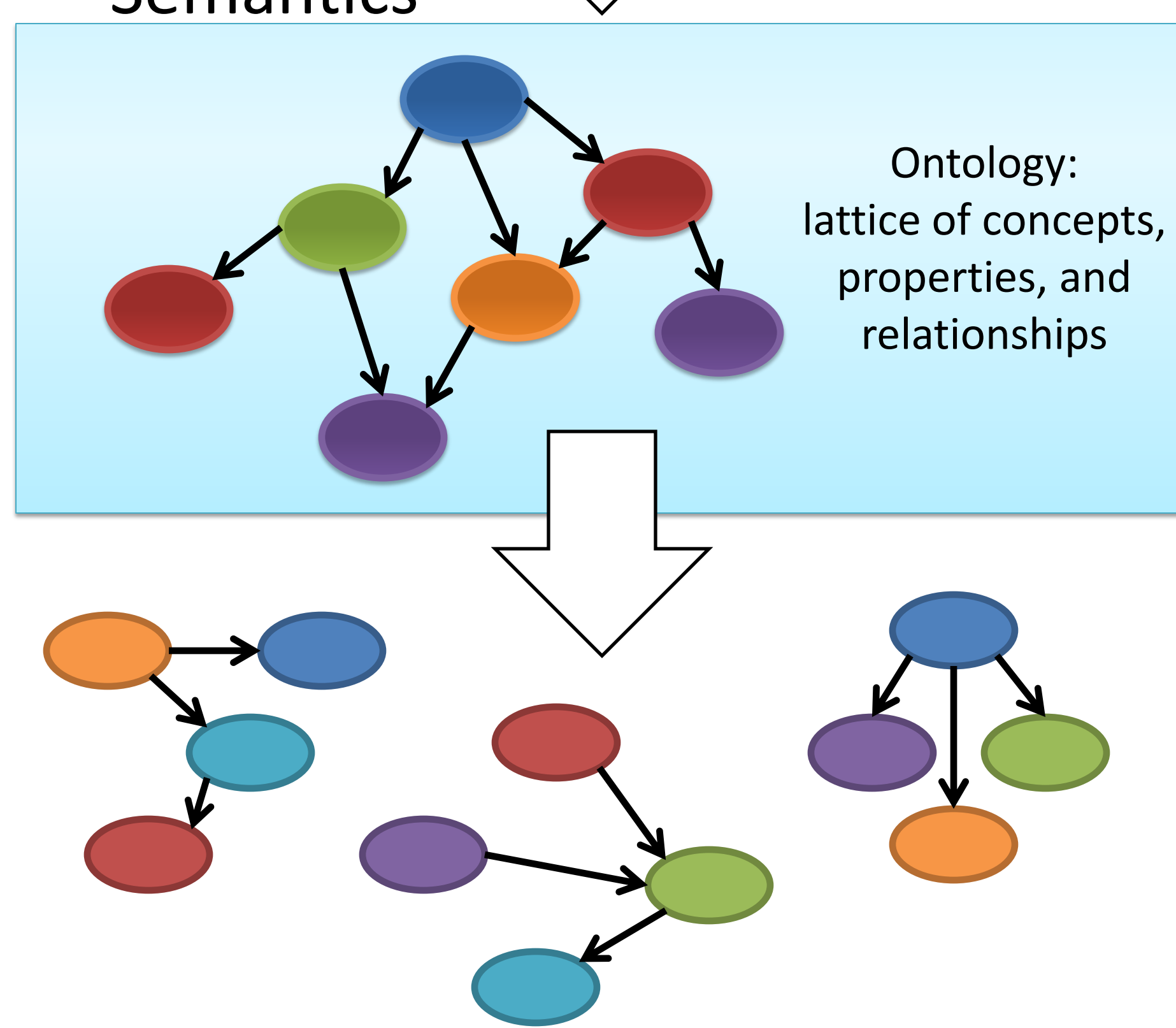
[2] Mikhail J. Atallah, Victor Raskin, jMichael Crogan, Christian Hempelmann, Florian Kerschbaum, Dina Mohamed, Sanket Naik, and Ira Moskowitz. 2001. Natural language watermarking: design, analysis, and a proof-of-concept implementation. In *Lecture Notes in Computer Science: Information Hiding (2137)*. Springer, Berlin, 185-200.

[3] Victor Raskin, Julia M. Taylor, and Christian F. Hempelmann. 2010. Ontological semantic technology for detecting insider threat and social engineering. In *Proceedings of the 2010 workshop on New security paradigms (NSPW '10)*. ACM, New York, NY, USA, 115-128. DOI=10.1145/1900546.1900563 <http://doi.acm.org/10.1145/1900546.1900563>

Ontological Semantic Technology processes natural language texts:

Example phrase: "control group features set"

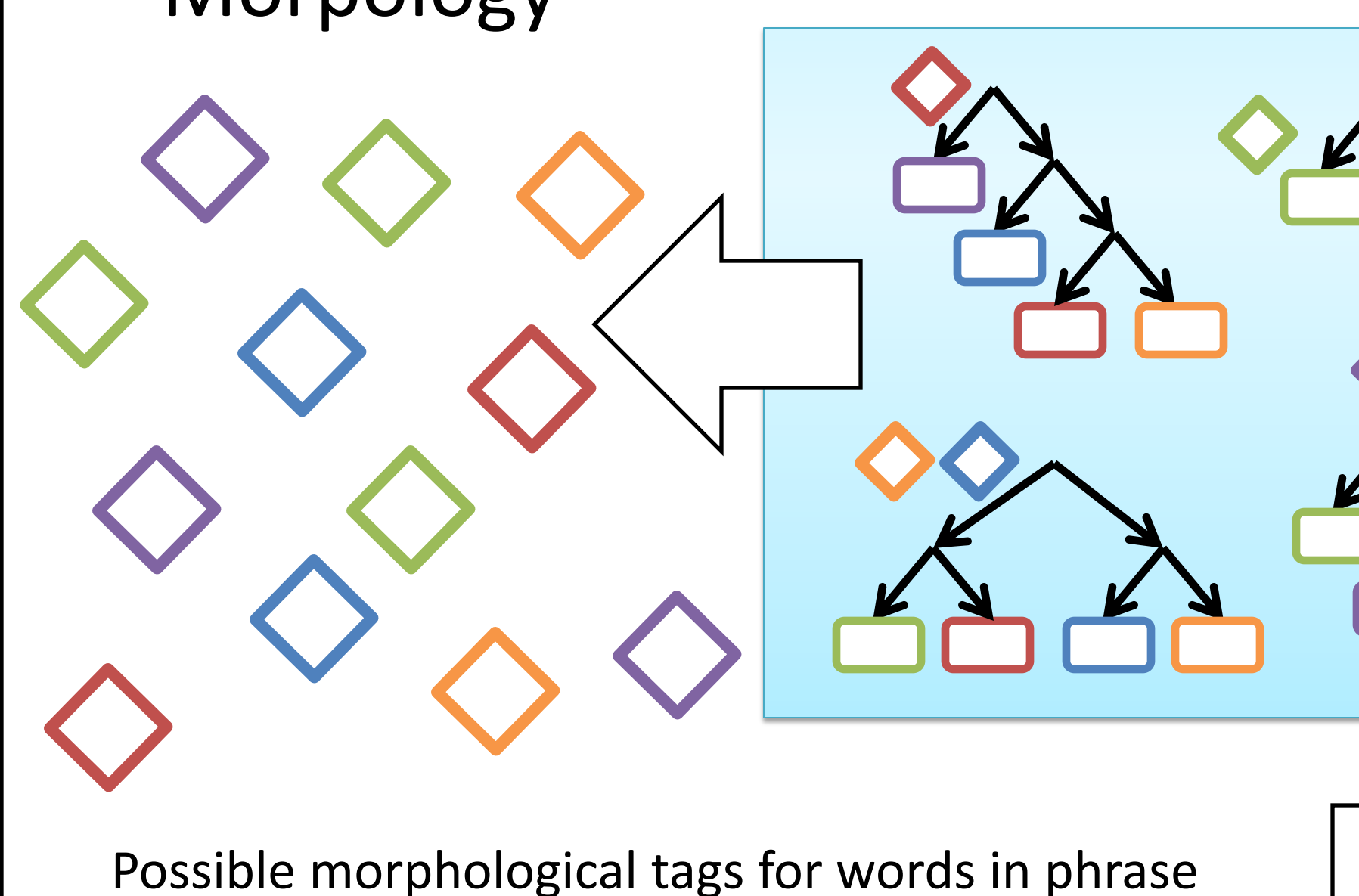
### Semantics



Possible interpretations:

- A group of controls features a set (of something)
- A group for the purpose of control features a set (of something)
- A command for someone/thing to control the set of features of some group
- A command for "control" to group some features which have been previously set

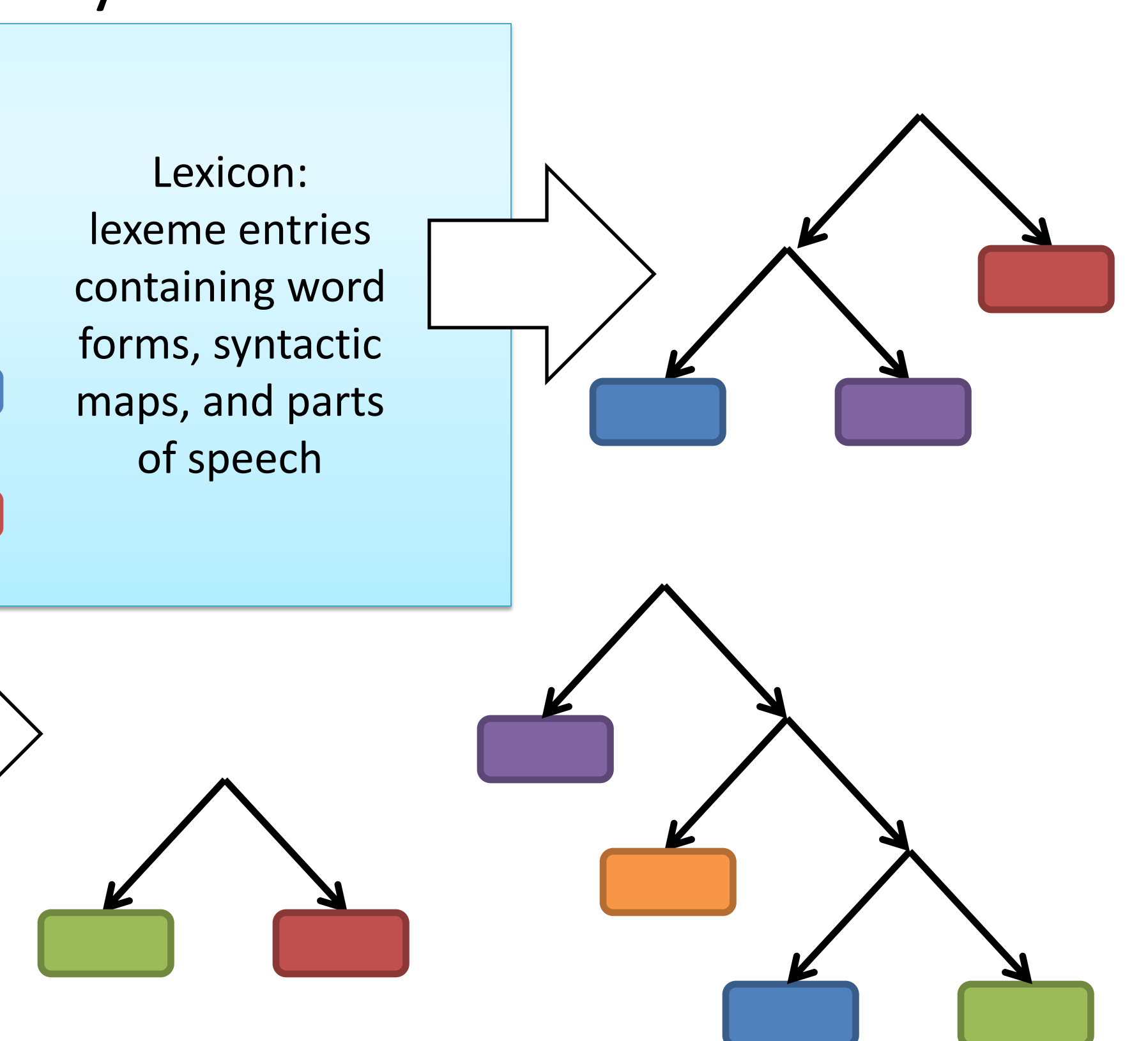
### Morphology



Possible tags:

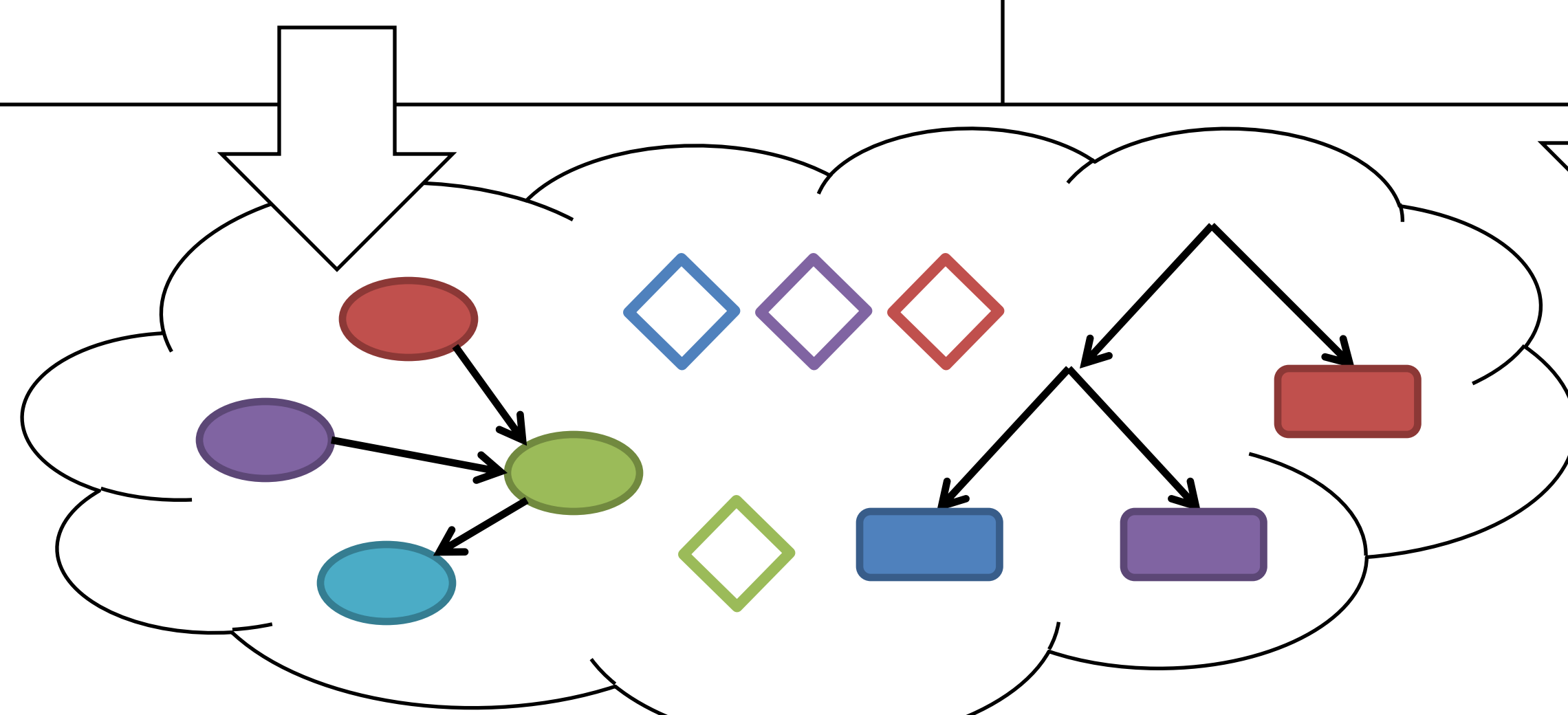
- "control"
  - 3<sup>rd</sup> person plural present verb, infinitive verb, singular noun
- "group"
  - 3<sup>rd</sup> person plural present verb, infinitive verb, singular noun
- "features"
  - 3<sup>rd</sup> person singular present verb, plural noun
- "set"
  - 3<sup>rd</sup> person plural present verb, infinitive verb, past verb, past participle, singular noun

### Syntax



Possible syntactic parses:

	"control"	"group"	"features"	"set"
1	noun	noun	verb	noun
2	verb	noun	noun	noun
3	verb	noun	noun	verb
4	name	verb	noun	noun



### Semantic, morphological, and syntactic characterization

- Structures can be modified for tamper-proofing or used to characterize an author.
- Interpretations can be reconciled with known information in a fact database.
- Reversing some components of this process can be used to paraphrase or translate texts into another language.

This research has been supported in part by NSF Research Grant #1012208, TC:Large:Collaborative Research:Anonymizing Textual Data and its Impact on Utility, to C. Clifton, PI; K. Chang, V. Raskin, L. Si, Co-PIs.