Rule-based and Similarity-based Operations in Artificial Grammar Learning

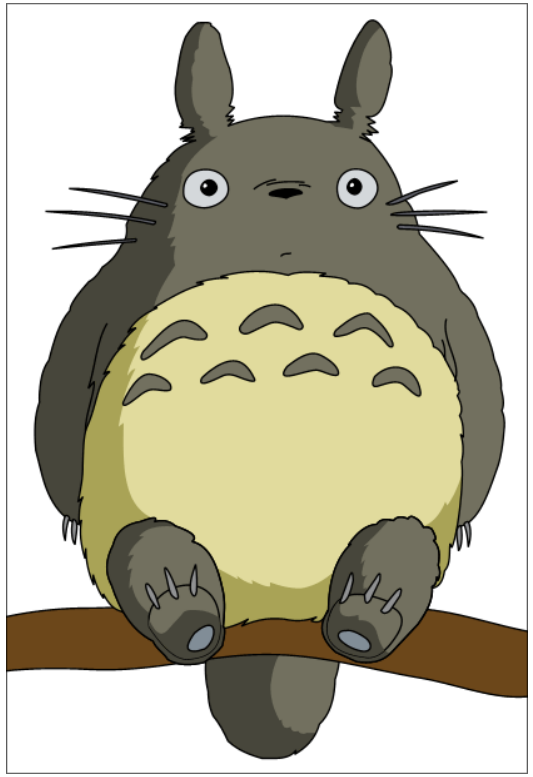
Zixian Chai

Daily Examples

- Rule-Based: Identify whether an integer is odd or even
- Similarity/Case-Based:
  - Apple: Vegetable? Fruit?
  - Tomato: Vegetable? Fruit? - Take longer time to think

Daily Examples – Rules? Similarity?

- Categorization: Identify the species of the Totoro
- Pattern Recognition: Identify a person by handwriting
- Reasoning: Analogical Reasoning
- Decision Making and Problem Solving: Encounter a complicated math problem and decide what theorems to use
- Language: Usage of pronouns, articles, propositions



(These sets of problems are overlapping)

Common Traits of These Examples

- Training Phase:

- A set of cases

- Follow certain rules

- ***Learned Information?***

- Test Phase:

- **Make judgements** on new cases

Artificial Grammar Learning Paradigm

Outline

- **Overview of rules and similarity**
- **Artificial Grammar Learning Paradigm (AGL)**
- **Research Design and Discussions**

Rules

- Production Rules: **if… then…(where…)** → Logical Inferences

Case	Input/Condition (>=1)	Output (>=1)	State (>=1)
1	RAINY and COLD	Do not go to play tennis	Healthy
2	SUNNY and COOL	Go to play tennis	Healthy
3	SUNNY and COOL	Do not go to play tennis	Sick

Similarity

- Relativeness between two cases

- Element Overlapping:

i.e. X {a,b,c} Y {b,c,d}

$$\text{Jaccard Similarity} = \frac{|X \cap Y|}{|X \cup Y|} = \frac{2}{4} = \frac{1}{2}$$

- Other methods: distribution, spatial similarity…

General Process

- Training Phase:

- A set of cases

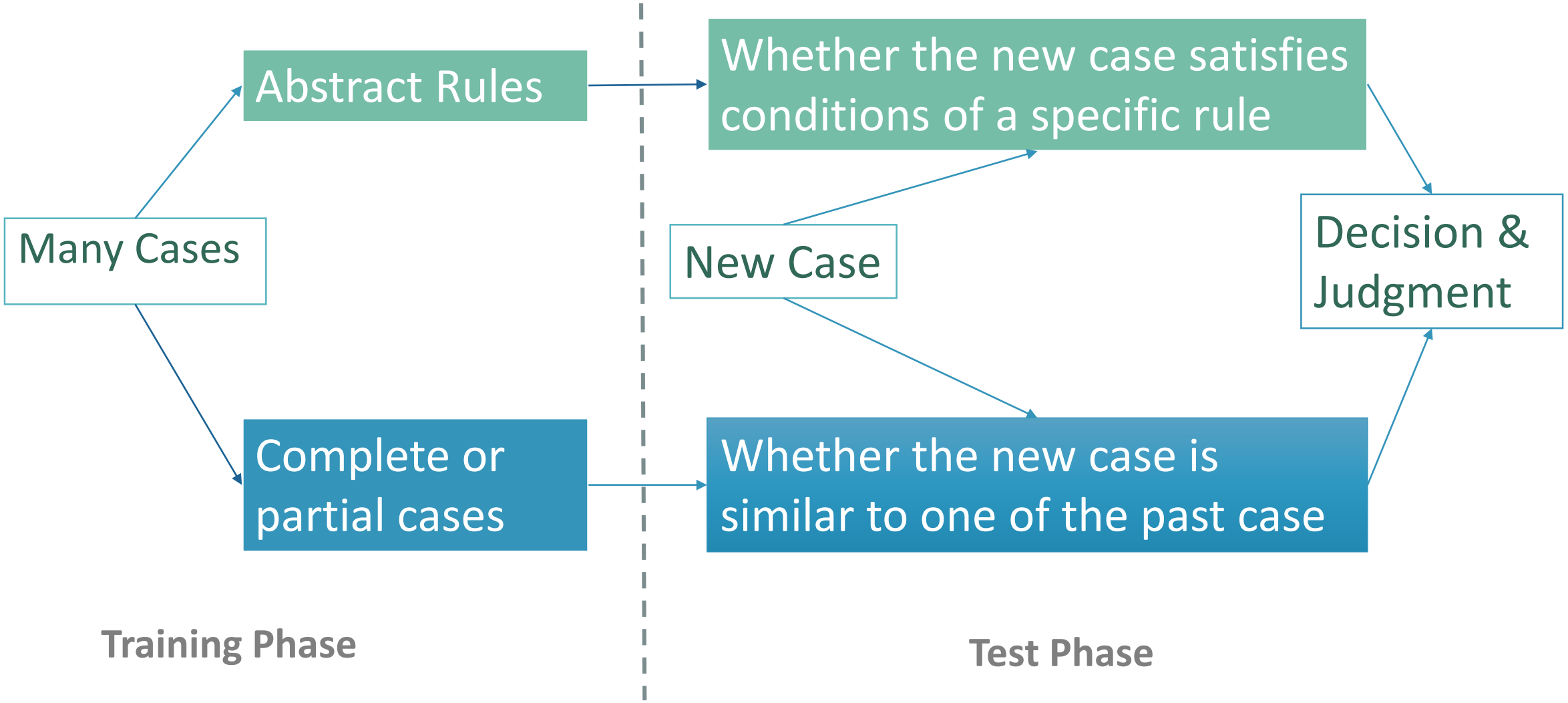
- Follow certain rules

- ***Learned Information?***

- Test Phase:

- **Make judgements** on new cases

Rule-based and Similarity-based Operations ?



Overlapping

Rule-based Operations involve Similarity Comparison

- Production Rules: Input will never be the same

Rules

- Production Rules: **if… then…(where…)** → Logical Inferences

Case	Input/Condition (Abstract)	Output	State
1	**RAINY (P 30%) and COLD (60F)**	Do not go to play tennis	Healthy
2	**SUNNY (P 3%) and COOL (75F)**	Go to play tennis	Healthy
new	**Precipitation 10% and 75F**	**?**	Healthy

A comparison between the new input and stored (abstract?) conditions
A comparison between the new state and stored (abstract?) states

Overlapping

Rule-based Operations involve Similarity Comparison:

- Production Rules: Input will never be the same: abstraction?

Similarity-based Operations involve Rules

- Comparison involves certain rules i.e.

$$\text{Jaccard Similarity} = \frac{|X \cap Y|}{|X \cup Y|}$$

- Old case -> Output1 and New case is similar to the old case

  Therefore, New Case -> Output1

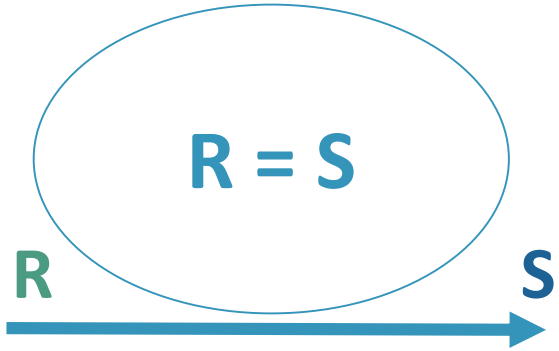
Debates: Connections? Distinctions?

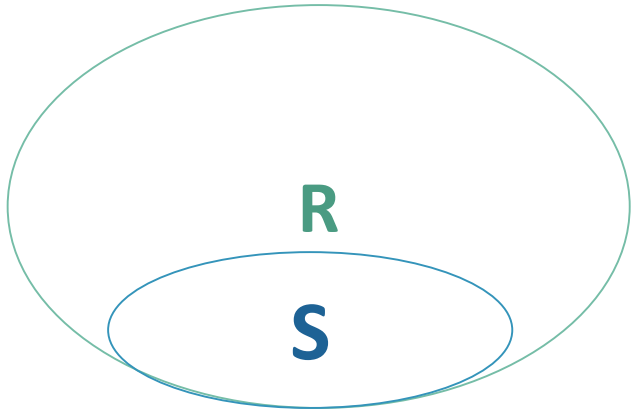
Rules and similarity are often regarded as two different concepts

Rules-Based Operations

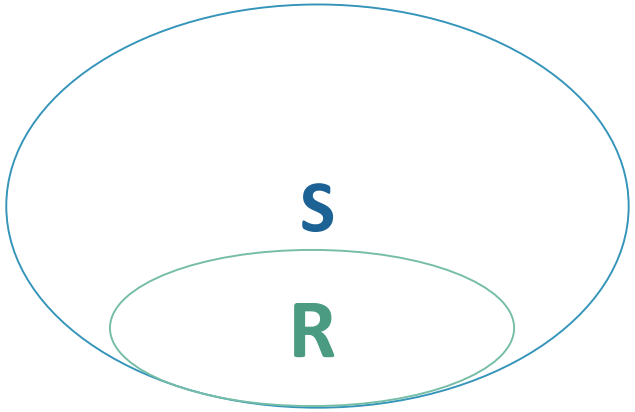
Similarity-Based Operations

Connections?



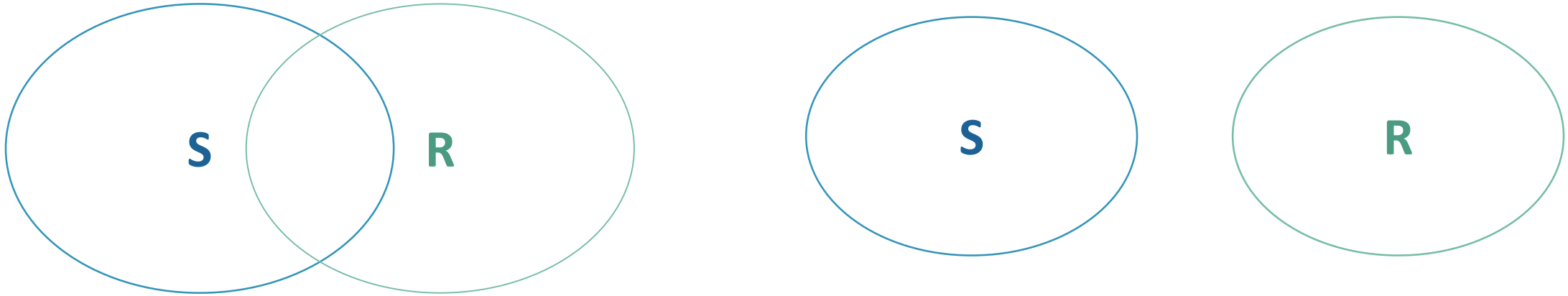


$S \subset R$



$R \subset S$

Distinctions?



Lack of Formal Definitions

Efforts of Clarification

- Abstraction of Representation

RBO - Abstraction/Generalization

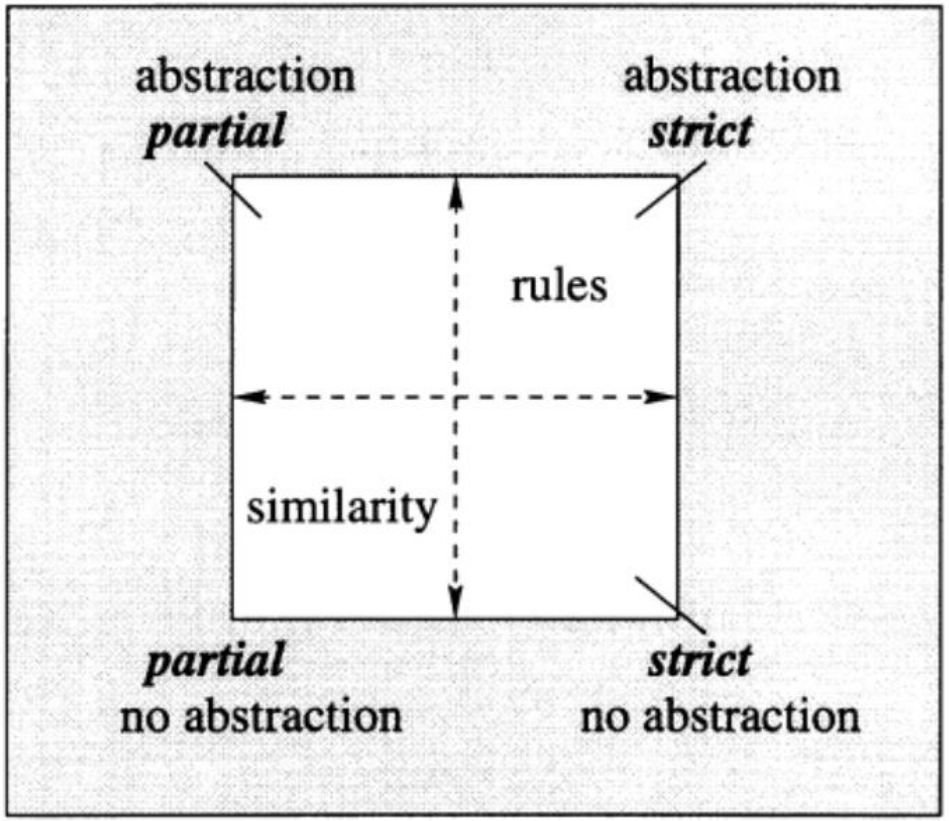
SBO - Specification

Perceptual Similarity and Analogical/Structural Similarity

SBO could also involve abstraction

Efforts of Clarification

- Hahn & Chater (1998), Abstraction (Relative) + Matching



representation -matching

non-representation -matching

- Strict: all or nothing
- Partial: Continuous

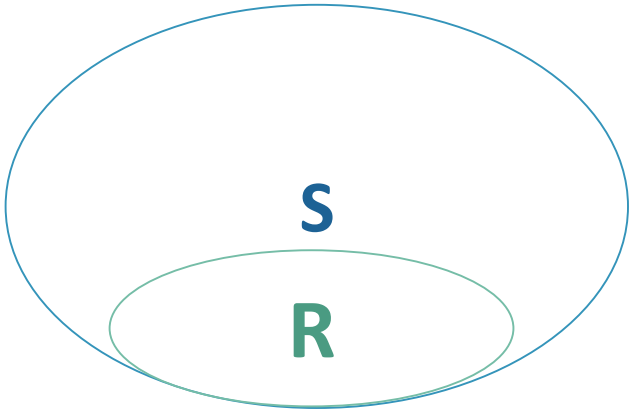
Efforts of Clarification

- Relative Abstraction

"Evidence of highly abstract mental representation is not evidence for rule-based processing and evidence for highly specific mental representation is not evidence for similarity-based processing."

----------Hahn & Chater, 1998, p202

Efforts of Clarification



$R \subset S$

- Pothos (2005), RBO as a specification of SBO

Categorization is determined by:

- A small subset of features that have higher weights - RBO

- All features have similar weights - SBO

Features could be abstract or specific.

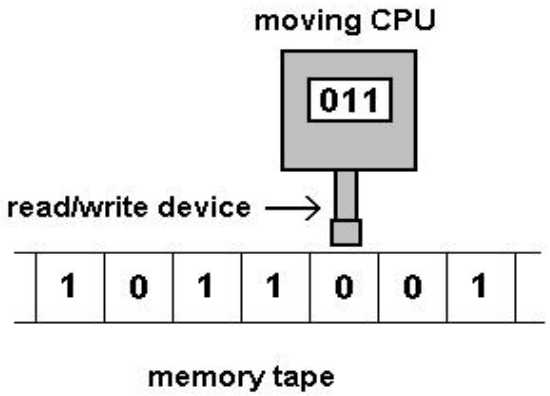
Efforts of Clarification

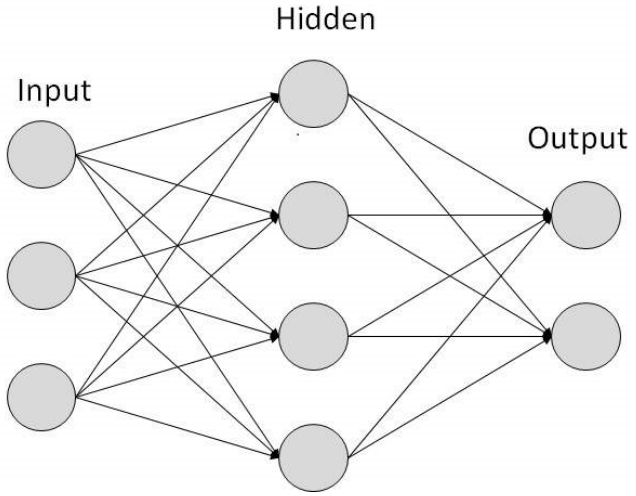
- RBO as a specification of SBO Numbers: Relativeness

Category α	Feature a	Feature b	Feature c	Operations
New Case 1	0.34	0.33	0.33	Similarity
New Case 2	0.99	0.01	0	Rule

Related Theories

- Associative Learning
- Prototypes
- Entropy: Quantification of uncertainty. People prefer certainty
- Information Processing Architectures: Turing Machine and Neural Network





General Process

- Training Phase:

- A set of cases

- Follow certain rules

- ***Learned Information?***

- Test Phase:

- **Make judgements** on new cases

- Avoidance of Previous Knowledge
- Simple Rules

Artificial Grammar Learning

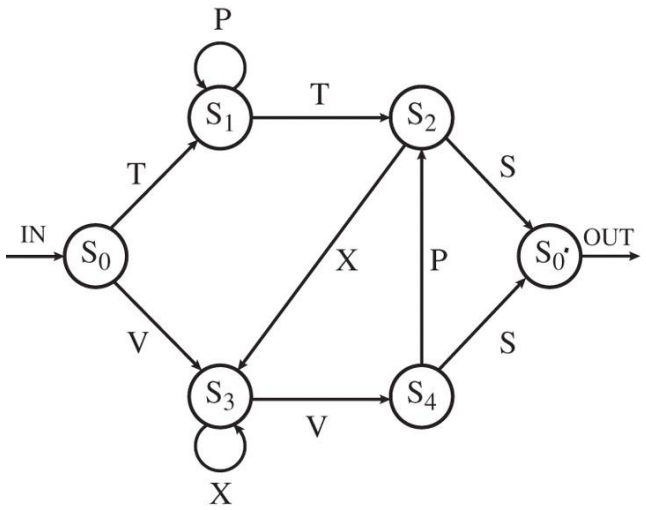
- A Finite State Grammar - Rules

a) An alphabet: Element can be specific or abstract {T, P, S, V, X}

i.e. {Red, Yellow, Green}, {Circle, Square, Triangle}, {Happiness, Sorrow, Anger}

b) A finite set of states, including a start state (S_0) and an end state (S_0')

c) A set of transition rules i.e. Letter S could transit from S



example strings:

TTS
TPPTS
TPPPTXVS
VXVS
VVS
TTXVPXVPS
TPPPPPPPTS
VXXXXXVS
TPPTXVPXXXXXVPXVPS

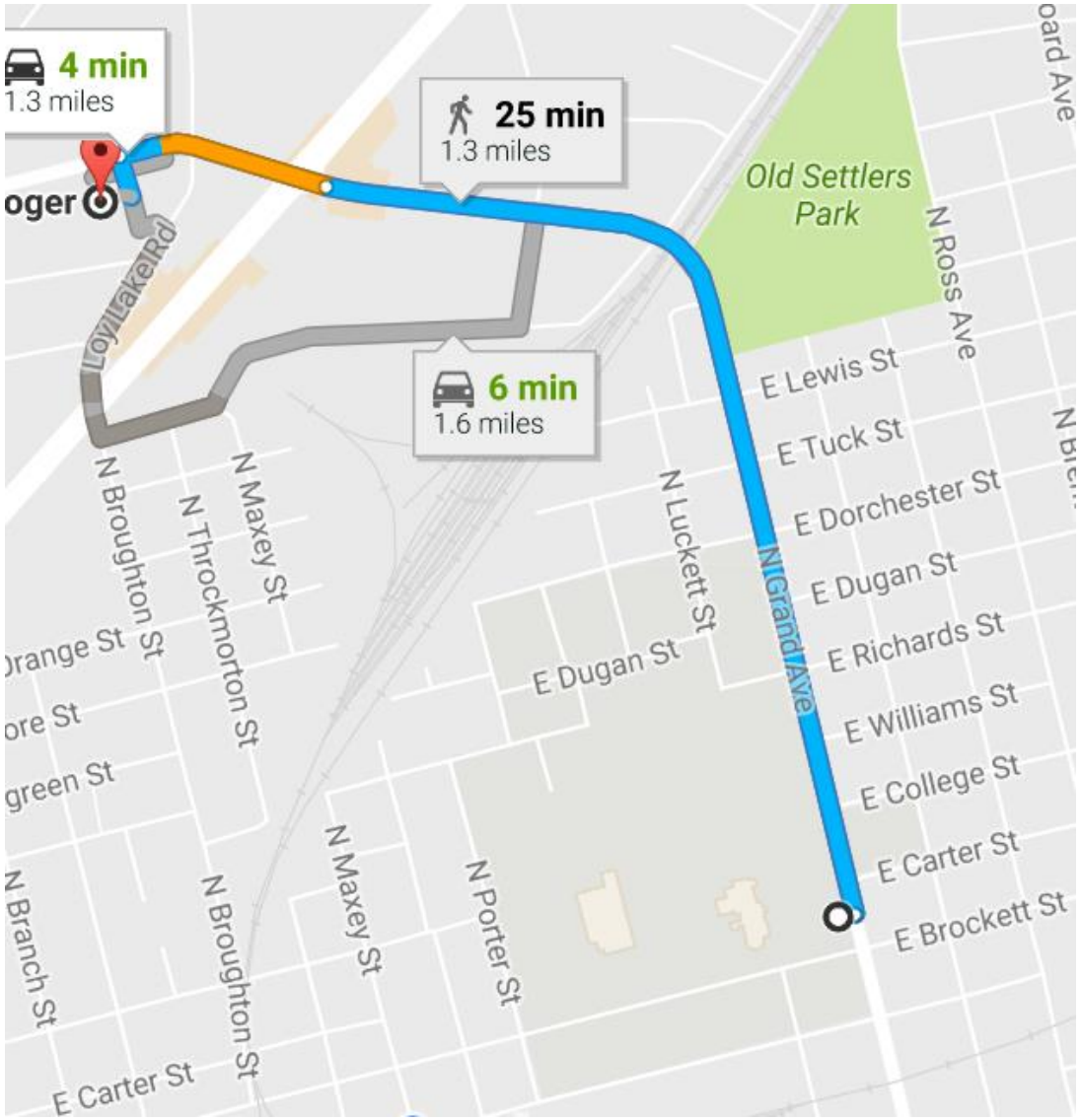
Finite State Grammar

- Open and Close the Door:

Alphabet: {Pull, Push}
States: {Open, Close}

- Game: level1 -> level2 -> level3

- Birth -> Infancy -> Childhood -> Adulthood -> Death



Artificial Grammar Learning

- Training Phase

Observe/memorize a set of items generated by the AG

- Test Phase

a. New items: X% Grammatical (G) and 1-x% Ungrammatical (UG)

b. Determine whether new items are G or UG

Artificial Grammar Learning

- Common Results

a. Above chance accuracy

b. Could not articulate how they make their judgments

Transfer Setting

- The training phase and the test phase use different alphabets
- Analogical Learning

	Letters	Modularity	Modularity
Training	{A, B, C, D, E}	Color	Musical Tones
Test	{O, P, Q, R, S}	Letters	Letters

Explanations

Rule-Related:

- Original grammar rules
- A set of correlated simple rules i.e. “A always follows B”, Head/Tail

Similarity-Related:

- Specific or average similarity
- Analogical similarity: MVXXZ *and* ABCCD
- Chunks i.e. AABC: AA, AB, BC, AAB, ABC

Entropy-Based: Degree of uncertainty

Experimental Questions

- How do rule-based operations and similarity-based operations interact in AGL?

Past research: Manipulation of either rules or similarity

- *Rules and Similarity vs. RBO and SBO -> Validity?*

- How do rules and similarity comparison involved in AGL?

Manipulation

Rule-Related Variable: The complexity of grammars

Similarity-Related Variable: Chunk Strength

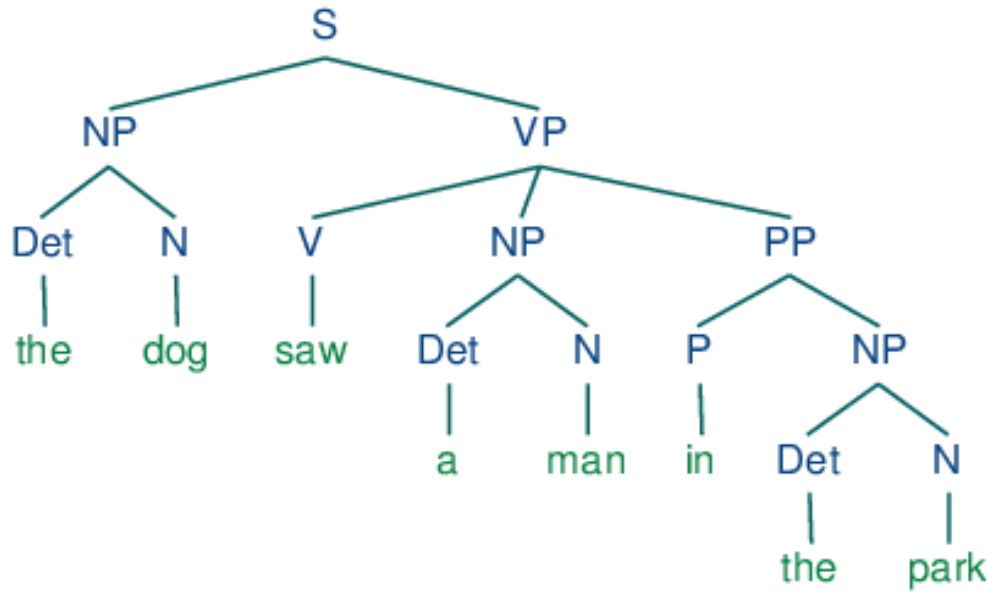
Standard and Transfer Experimental Settings

(Color and String)

Complexity of Grammars

- Finite State Grammar: Same level of abstraction

- Context Free Grammar: Hierarchical abstraction/representations



Chunk Strength

- Frequency of a chunk:

How many times a chunk appear previously in all training cases

- The averaged frequency of all two-letter and three-letter chunks in a specific new case

- Numerical, Continuous
- Correlated Grammar and Chunks
- Chunk strength balanced out specific similarity

- Low <4.5
- Medium 4.5-6.5 (inclusively)
- High >6.5

Experimental Design

	Finite State Grammar(FSG)			Context-Free Grammar(CFG)		
Condition	C1	C2	Control	C3	C4	Control
Training Phase	Letter Strings	Color Sequences		Letter Strings	Color Sequences	
Test Phase	Letter Strings + Color Sequences			Letter Strings + Color Sequences		

Test Phase: Chunk Strength - High, Medium, Low

Human Performance Expectation

	Significant	Insignificant
Main effect of RV	RV is involved	RV is not involved
Main effect of SV	SV is involved	SV is not involved
Interaction Effect	RV and SV interact with each other	?

RV: Rule-related Variable SV: Similarity-related Variable

Computational Models

- Simple Recurrent Network: Information is distributed throughout the network
- Competitive Chunking Network: Information is stored locally

- Comparison: Pearson's r Correlation, Kendall's τ Ranking Correlation

Limitations

- *Rules and Similarity vs. RBO and SBO*

- *Statistical Power: 2 (between-group) x 3 (within-group) = 6 conditions + transfer + control*

- *Possible Confounding Variables*

Controlled:

- Specific Similarity: >4

Not Controlled:

- Averaged Similarity: the bias factor, Entropy, Head/Tail Effect (partially controlled), Different types of violations in ungrammatical items…

Limitations

- Control Group

	Training Phase	Test Phase	Limitations
Control 1		Same Test Phase	Confusion on judgments
Control 2	Random Strings	Same Test Phase	Influence of the training phase

Something Else: Psychology and Big Data

- Training Phase - Data

a. Parameter Weighting/Discovery

b. Structure Weighting/Discovery

- Test Phase

Make prediction/judgment on new data

A Large Data Set – Increased Power

Something Else: Psychology and Big Data

- Practices

- Searching History and Relativeness Recommendation
- Customer Relationship Management, Employee Performance Analysis
- Semantic Analysis
- Social Medias Data Analysis
- Brain Image Analysis
- Educational Software
- Games
- Self-driving Car

Something Else: Psychology and Big Data

- Available Public Data Sets
- Available Analysis Software
- Development of Easy-to-use Programming Languages

- Research: Ethics ?
- Practice: Information Security?

Reference

- Hahn U. & Chater N. (1998). Similarity and rules: distinct? Exhaustive? Empirically distinguishable? *Cognition, 65*, 197-230.
- Jamieson R. K., Nevzorova U., Lee G. & Mewhort D.J.K.(2016). Information theory and artificial grammar learning: Inferring grammaticality from redundancy. *Psychological Research, 80*, 195-211
- Kinder A. & Lotz A. (2009) Connectionist models of artificial grammar learning: what type of knowledge is acquired? *Psychological research: 73,* 659-673.
- Pothos, E.M. (2005). The rules and similarity distinction. *Behavioral and Brain Sciences,* 28, 1-49.
- Pothos, E.M. (2007) Theories of artificial grammar learning. *Psychological Bulletin, 133*(2), 227-244.
- Reber A. S. (1967). Implicit learning of artificial grammar. *Journal of verbal learning and verbal behavior, 6*, 855-863.
- Reber A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of experimental psychology, 81*(10), 115-119.