Statistical learning mechanisms: Pearl 2021 Chomsky handbook 1.2.1 + Pearl 2021 JoCL UG+Stats 2.2

CompCog Models of LangAcq

Kellen Glass

angAcq

Contents

Counts/Statistical Learning	3
Reinforcement Learning	4
Tolerance/Sufficiency Principles	5
Bayesian Inference/Modeling	7
UG	8
UG + Statistical Learning	9
Effects on Systems	11
Conclusion	12



Counts and Statistical Learning

- Statistical learning
 - Has to do with counting
 - The learning component of acquisition (inference)
 - A child has to know what to count
- We use counts to create probabilities
 - Seeing something 3 times out of 10 = .30 probability
- Inference and counting go hand in hand
 - A child can do inference over a set of lexical items

Reinforcement Learning

- Involves internal work a child is doing to understand a language
- Good/proper behavior is encouraged
 - When there is +wh-movement, that option is reinforced, increasing the likelihood that it is repeated
- There is evidence that very young children can use this
 - As young as 10 weeks

The Tolerance and Sufficiency Principles

- Tolerance Principle
 - How many exceptions to an internalized rule we tolerate before we get rid of the rule
- Sufficiency Principle
 - How many rule following items is enough to warrant the maintenance of a rule
- Inference mechanism that operates over certain counts
 - Leads to creation of principles which leads to rule generalization
 - A pattern used for a set of items

The Tolerance and Sufficiency Principles cont.

- Helps children learn
 - Verb forms, syntactic positions, default stress patterns
- Knowledge storage and retrieval
 - Ranked by frequency
 - Fastest to retrieve information when it is related
 - If something is too different, the average retrieval time gets too long
- Used to understand exceptions
 - Drink -> (past) drank, not drinked
 - Understand syntactic position changes and different stress patterns too
- Not as much evidence of this in younger children
 - Mostly 5-8 year olds

Bayesian Inference/Modeling

- Operates over probabilities
- Bayesian Model
 - \circ H = Hypothesis
 - \circ **D** = **D**ata
 - P(h|D) = probability of each hypothesis
 - $P(h|D) \propto P(D|h) * P(h) =$ likelihood of data given hypothesis and prior hypothesis
 - Capture how compatible a model is with data
 - Prior hypothesis shows how plausible the current one is
- Young children are likely capable of Bayesian inference

UG (Universal Grammar)

- Humans have a natural understanding of grammar
- Defines what linguistic hypothesis to consider
 - Acts as parameter for
 - constructions/considerations of
 - language/hypotheses about a language's system (word order, inflection, movement)

UG + Statistical Learning

- Statistical Learning compliments UG
 - UG can be used to correctly identify a hypothesis/parameters for a language
 - UG identifies what to count, statistical learning works on creating probability
 - Statistical learning mechanism does not change once a child has counts of a relevant unit, allows for inference

UG + Statistical Learning cont.

Table 1: Common inference mechanisms in statistical learning that are used by UG+stats proposals for different morphology and syntax phenomena: basic syntactic categories (**syn cat**), basic word order (**word order**), inflectional morphology (**infl mor**), showing a temporary lack of inflection (**no infl**), movement (**mvmt**), and constraints on utterance form and interpretation (**constr**).

· · · · · · · · · · · · · · · · · · ·	syn cat	word order	infl mor	no infl	mvmt	constr
Basic counts & probabilities	~	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Bayesian inference	~		√		\checkmark	\checkmark
Reinforcement learning		\checkmark		\checkmark	\checkmark	
Tolerance & Sufficiency			\checkmark			

• What inference mechanisms are used by particular UG+stats proposals

Effects on Systems

- Bayesian Modeling
 - UG defines parameters to determine if a language has a certain type of grammatical system
 - +/- wh-movement
- Reinforcement Learning
 - Linear reward-penalty system
 - 2 choices, one is rewarded the other is penalized

Conclusion

- Statistical Learning
 - Counts/probability
- Reinforcement Learning
 - Internally encourage proper behavior (operant conditioning)
- Tolerance/Sufficiency Principles
 - Rule generalization/maintenance

Conclusion cont.

- Bayesian Inference/Modeling
 - Proving the relevance of a hypothesis with data
- UG
 - Defines parameters
 - Identifies hypothesis
- UG + statistical learning
 - UG identifies what to count
 - Statistical learning finds probability

Thank You

Class Name

Student Name