Using Salience and Hypothesis Evaluation to Learn Object Names in Real Time

Jon Stevens
University of Pennsylvania

Word learning is a hard nut to crack, but recent experimental and computational research has put us in a position to simplify the problem. Following up on modeling research done by Yu and Ballard (2007) and Frank, Goodman, and Tenenbaum (2009), as well as experimental work done by Medina, Trueswell, Snedeker, and Gleitman (2009), we propose an online computational model of object name learning which is able to learn a precise and reasonably thorough lexicon from a small amount (<20 minutes' worth) of real data. This model represents a reduction in complexity from previous models and a step in the direction of psychological plausibility.

Extending Yang's (2002) approach to syntactic and morphological acquisition, we model early word learning as a process of updating probabilities within a constrained hypothesis space. We also utilize independently motivated external cues. Following Yu and Ballard (2007), we integrate gestural and prosodic cues into the learning model, which we believe come at little conceptual cost given infants' sensitivity to visual and auditory salience and to the actions of others (see Bloom 2000 for a summary). This allows simple probabilistic reward-penalty algorithms to operate in real time yielding satisfactory results with minimal computation.

The algorithm (Fig. 1) accesses and updates a lexicon, which represents the learner's beliefs about word-object pairings at any given point in time. Upon hearing a word, the learner checks the lexicon for a hypothesis. If one exists, then that hypothesis is evaluated against the current situation, and its probability is updated using the Linear Reward-Penalty (LR-P) scheme (Bush and Mosteller 1951, Yang 2002). In the absence of a hypothesis, the learner will resort to multiple-candidate cross-situational learning; the learner will begin with the assumption that a novel word does not refer and then reward *all* candidate referents for each utterance of that word. During this multiple-candidate rewarding process, mappings between words with sentential stress and objects that are being gestured to in an obvious way receive a larger probability boost. A word-object pairing is added to the lexicon if and only if that pairing's probability reaches a given threshold value, and removed if it drops below that threshold. This model outperforms a strictly multiple-candidate model, which does not utilize hypothesis evaluation, and it finds an empirical basis in recent eye-tracking studies (Medina et al 2009) that show that learners test particular beliefs about word meanings rather than always attending to all possible meanings.

Simulations were run on a hand-coding of two videos of mother-child interaction from the CHILDES database (MacWhinney 2000) and evaluated using the average F-score (harmonic mean of precision and recall) of the output lexicon for each model against a gold standard. Fig. 3 compares the performance of the online learning models with implementations of the Bayesian batch learning algorithm presented in Frank et al (2009). The lexicon most commonly learned by the best online model is given in Fig. 4. These results show that a simple and empirically motivated name learning algorithm can produce satisfactory results from real data, extending current conceptions of linguistic learning to the semantic domain, and creating a framework that may be extensible to semantic learning processes beyond object name learning.

References

Bloom, Paul. 2000. How Children Learn the Meanings of Words. Cambridge, MA: MIT Press. / Bush, Robert and Frederick Mosteller. 1951. A Mathematical Model for Simple Learning. Psychological Review 68, 313-23. / Frank, Michael, Noah Goodman, and Josh Tenenbaum. 2009. Using Speakers' Referential Intentions to Model Early Cross-Situational Word Learning. Psychological Science 20:5. / MacWhinney, Brian. 2000. The CHILDES Project: Tools for Analyzing Talk. Vol. 2: The Database (3rd ed.). Mahwah, NJ: Erlbaum. / Medina, Tamara Nicol, John Trueswell, Jessie Snedeker, and Lila Gleitman. 2009. Rapid Word Learning under Realistic Learning Conditions. Paper presented at LSA Annual Meeting, San Francisco, CA, Jan. 9. / Yang, Charles. 2002. Knowledge and Learning in Natural Language. Oxford: Oxford University Press. / Yu, Chen and Dana Ballard. 2007. A Unified Model of Early Word Learning: Integrating Statistical and Social Cues. Neurocomputing 70, 2149-2165.

	Given a possibly referential word *w*, a randomly-ordered set of possible object referents *O*, and a set of word-object pairs constituting a lexicon *L*:
HYPOTHESIS INTRODUCTION	If *w* is a novel word, assume with probability 1 that it does not refer to an object. For each *o* in *O*: If *o* has no probability value associated with *w*, assign to *o* a probability of (1 / n), where n is the number of candidate meanings for *w*. Renormalize the probability vector for *w*.

PROBABILITY UPDATING	If *w* is an element of $\cup L$:
	    If *w*'s hypothesized meaning *h* is an element of *O*, reward *h*.
	    Else, penalize *h*.
	If *w* is not an element of $\cup L$:
	    For each *o* in *O*:
	        If *o* is not an element of $\cup L$, reward *o*.

LEXICON CONSTRUCTION	After each utterance, update *L* to include all and only those word-object pairings that have an associated probability greater than some threshold value τ.

Fig. 1: An Online Name Learning Algorithm with Hypothesis Evaluation

reward *h*	$p(h) = p(h) + \gamma\,(1 - p(h))$ For all $h' \neq h$: $p(h') = p(h') * (1 - \gamma)$	penalize *h*	$p(h) = p(h) * (1 - \gamma)$ For all $h' \neq h$: $p(h') = (\gamma / n\text{-}1) + p(h') * (1 - \gamma)$

Fig. 2: Linear Reward-Penalty Functions (where γ is some constant between 0 and 1)

	Precision	Recall	F-score
Bayesian (Frank et al 2009)	0.36	0.29	0.32
Bayesian w/ stress and gesture	0.72	0.38	0.52
Online probability updating, strictly multiple-candidate	0.24	0.06	0.10
Online w/ hypothesis evaluation	0.36	0.06	0.10
Online w/ hypothesis evaluation, stress, and gesture	0.92	0.32	0.48

Fig. 3: Model Comparison

Word	Object	Word	Object	Word	Object	Word	Object
book	**book**	**kittycat**	**cat**	**piggies**	**pig**	meow	cat
bear	**bear**	**bigbird**	**bird**	**hat**	**hat**	**sheep**	**sheep**
bunny	**bunny**	**ring**	**ring**	**moocow**	**cow**	**bird**	**duck**

Fig. 4: Most Common Lexicon Output by Online Model with Salience Cues and Hypothesis Evaluation