



++

# **Session 2. Strategy 1: Simultaneously Inferring Phrases and Topics**

# Strategy 1: Simultaneously Inferring Phrases and Topics

---

- **Bigram Topic Model [Wallach'06]**

- Probabilistic generative model that conditions on previous word and topic when drawing next word

- **Topical N-Grams (TNG) [Wang, et al.'07]**

- Probabilistic model that generates words in textual order
- Create n-grams by concatenating successive bigrams (a generalization of Bigram Topic Model)

- **Phrase-Discovering LDA (PDLDA) [Lindsey, et al.'12]**

- Viewing each sentence as a time-series of words, PDLDA posits that the generative parameter (topic) changes periodically
- Each word is drawn based on previous m words (context) and current phrase topic

- High model complexity: Tends to overfitting; High inference cost: Slow

# TNG: Experiments on Research Papers

Reinforcement Learning			Human Receptive System		
LDA	<i>n</i> -gram (2+)	<i>n</i> -gram (1)	LDA	<i>n</i> -gram (2+)	<i>n</i> -gram (1)
state	reinforcement learning	action	motion	receptive field	motion
learning	optimal policy	policy	visual	spatial frequency	spatial
policy	dynamic programming	reinforcement	field	temporal frequency	visual
action	optimal control	states	position	visual motion	receptive
reinforcement	function approximator	actions	figure	motion energy	response
states	prioritized sweeping	function	direction	tuning curves	direction
time	finite-state controller	optimal	fields	horizontal cells	cells
optimal	learning system	learning	eye	motion detection	figure
actions	reinforcement learning rl	reward	location	preferred direction	stimulus
function	function approximators	control	retina	visual processing	velocity
algorithm	markov decision problems	agent	receptive	area mt	contrast
reward	markov decision processes	q-learning	velocity	visual cortex	tuning
step	local search	goal	vision	light intensity	moving
dynamic	state-action pair	space	moving	directional selectivity	model
control	markov decision process	step	system	high contrast	temporal
sutton	belief states	environment	flow	motion detectors	responses
rl	stochastic policy	system	edge	spatial phase	orientation
decision	action selection	problem	center	moving stimuli	light
algorithms	upright position	steps	light	decision strategy	stimuli
agent	reinforcement learning methods	transition	local	visual stimuli	cell



# TNG: Experiments on Research Papers (2)

Speech Recognition			Support Vector Machines		
LDA	$n$ -gram (2+)	$n$ -gram (1)	LDA	$n$ -gram (2+)	$n$ -gram (1)
recognition	speech recognition	speech	kernel	support vectors	kernel
system	training data	word	linear	test error	training
word	neural network	training	vector	support vector machines	support
face	error rates	system	support	training error	margin
context	neural net	recognition	set	feature space	svm
character	hidden markov model	hmm	nonlinear	training examples	solution
hmm	feature vectors	speaker	data	decision function	kernels
based	continuous speech	performance	algorithm	cost functions	regularization
frame	training procedure	phoneme	space	test inputs	adaboost
segmentation	continuous speech recognition	acoustic	pca	kkt conditions	test
training	gamma filter	words	function	leave-one-out procedure	data
characters	hidden control	context	problem	soft margin	generalization
set	speech production	systems	margin	bayesian transduction	examples
probabilities	neural nets	frame	vectors	training patterns	cost
features	input representation	trained	solution	training points	convex
faces	output layers	sequence	training	maximum margin	algorithm
words	training algorithm	phonetic	svm	strictly convex	working
frames	test set	speakers	kernels	regularization operators	feature
database	speech frames	mlp	matrix	base classifiers	sv
mlp	speaker dependent	hybrid	machines	convex optimization	functions