

# Prelim Notes

Dylan

Eric

Robert

Shubham

## 1. Natural Language Processing and Speech

### 1.1. Chapter 1-4 (July 18)

#### Discussions

- Are n-grams defined over word forms or word roots ?  
Ans - application specific.
- Good-Turing smoothing : we derived Eq 4.27 using Eq 4.26 and briefly discussed the approximation used for larger N (Simple-Good turing)
- Brief discussion of interpolation and katz-backoff.
- Discussed back-off in Kneser Ney. An unanswered question was regarding implementation of back-off from n-grams to (n-1)-grams for  $n > 1$  (do we use context or back-off to Kneser-Ney probabilities ?)

#### Topics for Review

- Kneser-Ney
- Perplexity
- Good Turing (formula)

### 1.2. Chapter 5-9 (July 22)

#### Discussions

- Maximum entropy Markov model (how this differs from hidden Markov model)
- Hidden Markov model speech recognition (how the model works)

#### Topics for Review

- Part of speech tagging (hidden Markov models and rule based)
- Maximum entropy and maximum entropy markov models
- features for speech (pitch, amplitude, spectrograms, formants)
- Cepstrum
- Speech recognition with hidden Markov model

### 1.3. Chapter 10, 12-13 (July 25)

#### Discussions

- Viterbi Approximation
- Only simple language models
- multi-pass – do coarse solution proposal, exact comparison of proposals
- word lattices – not 100% on how to build it, but its probably just a modified viterbi
- confidence
- A\* decoding – fast match;  $g(\text{cost}) + \text{partial path}$  is confusing
- triphones – bigram to trigram model; state space expansion; decision tree
- MMIE – wtf
- adaptation – gender; Maximum Likelihood Linear Regression (MLLR)
- Dependency vs CFG
- Bottom-up vs top-down
- parsing hard b/c ambiguity
- CKY

#### Topics for Review

- Stack/ A\* decoding
- Multi-Pass
- Tri-Phones
- Adaptation
- Know Context Free Grammars
- Chomsky Normal Form

## 1.4. Chapter 14-15, 17 (July 29)

### Discussions

- Inside-outside algorithm relation to EM? (It seems to be closer to forward-backward)

### Topics for Review

- PCFGs, probabilistic CKY, inside-outside algorithm
- Mitigating poor independence assumptions: splitting tags, split and merge
- Mitigating lack of lexical conditioning: head word tagging, Collins parser
- Feature structures - used to prevent an explosion of grammar rules, augmented Earley parser

## 1.5. Chapter 18, 20-21 (August 1)

### Discussions

- Quantifier scope ambiguity
- WSD algorithms overview - feature based / dictionary-context similarity and selectional restrictions
- Distributional and Thesaurus methods for word similarity
- Coreference resolution algorithms

### Topics for Review

- Understanding semantic attachments to syntax.
- Various ways of under-specification to deal with ambiguity.
- Semantic Role labelling (book wasn't clear so maybe look at Dan's slides)
- Pronominal anaphora resolution
- Coreference resolution algorithms

## 1.6. Chapter 23-25 (August 5)

### Discussions

- In SMT, why do we need to use Bayes' rule
- Make sure we understand how phrase-based stack decoding works

### Topics for Review

- tf-idf
- query answering
- document summarization, featurize sentences, discourse parsing
- sentence simplification
- conversational agents with MDPs
- Vauquois triangle, syntactic transfer, direct translation
- statistical machine translation, noisy channel model
- translation models
- alignment in MT, HMM alignment model
- phrase-based stack decoding
- log-linear model for MT

## 2. Artificial Intelligence: A Modern Approach

### 2.1. Chapter 1-5 (July 22)

#### Discussions

- variants of  $A^*$  (limited memory, weighted,  $LRTA^*$ )
- And-or search trees (their different uses)
- alpha-beta pruning (work through an example)

#### Topics for Review

- BFS, DFS,  $A^*$ , limited-memory  $A^*$
- Iterative deepening DFS
- Heuristics for  $A^*$
- Simulated annealing, evolutionary algorithms
- And-or search trees
- $LRTA^*$
- Minimax
- Alpha-beta pruning
- Expectimax

### 2.2. Chapter 6-8 (July 25)

#### Discussions

- K-consistency – generalization of individual consistency
- Reduction to binary constraints – introduce variables, dual graphs
- path consistency – remove pairs or singletons
- backjumping – completeness; backjumping undoes any variables that were set after the node you back-jump to
- Forward Checking  $\subset$  Arc Consistency
- 6.5 – favorite section for questions
  - Tree Structured CSP – enforce directed arc consistency  $\rightarrow$  no backtracking needed
  - Cut sets – iterate over all possible assignments to the cut set; cut set leaves remaining graph as tree
  - Tree decomposition – junction tree for CSPs! – graph with subsets of variables, tree structured solver enforces variable consistency.
- Basic logic modus ponens/and resolution

- models  $\rightarrow$  possible worlds, simplest inference = model-checking –  $\alpha \implies \beta$  iff  $M(\alpha) \subseteq M(\beta)$ .
- Forward Chaining – Horn Clauses + Definite Clauses; important idea is the lhs of implications are positive conjunctions. That property saves backtracking over incorrect assignments.
- DPLL – DFS with Heuristics; what do we need to know
- Satisfiability Problems are usually easy
- frame problem, fluents,
- belief states – represent 'belief' as things that could be true
- SATPlan – be super verbose or problems arise
- object/relations/functions
- quantifiers  $\exists, \forall$

#### Topics for Review

- CSPs – know them
- consistency: node, arc, path, strong-k
- heuristics: minimum remaining value, degree, least constrained
- backjumping/conflict directed
- local search
- Tree Structured- CSPs
- knowledge base
- syntax
- possible worlds/models
- entailment
- horn clauses/forward chaining
- (logical) state estimation

### 2.3. Chapter 9-12 (July 29)

#### Discussions

- Propositionalization of first-order logic statements
- Generalized Modus Ponens, unification
- Forward chaining in FOL - NP hard, can use heuristics and incremental forward chaining
- Backward chaining

- Resolution - conversion to CNF (including Skolem functions), refutation-completeness
- Equality in FOL - demodulation, paramodulation
- PDDL - states, actions
- Planning as a search problem, heuristics
- Planning graphs, heuristics, GRAPHPLAN
- Partially ordered plans
- Critical path method, heuristics
- Hierarchical planning, high-level actions, angelic semantics
- Nondeterminism in planning - percept augmentation to PDDL, belief states as conjunctions of positives, replanning
- Multiagent planning

#### Topics for Review

- Inference in FOL - Generalized Modus Ponens, unification, forward/backward chaining, resolution
- PDDL, planning problems, search strategies, planning graphs
- Partially ordered plans
- Scheduling - critical path method, hierarchical planning, nondeterminism in planning
- Learning using relevance information, Inductive logic programming

#### 2.4. Chapter 17, 19, 25-27 (August 1)

##### Discussions

- Value Iteration for POMDPs (is it really used ?)
- EBL and its uses

##### Topics for Review

- Bellman Equation (know how to write it)
- Review ways to solve POMDPs
- EKF, SLAM (from Pieter's slides)

### 3. Reinforcement Learning

#### 3.1. Chapter 1-5 (August 5)

##### Discussions

- No discussion

##### Topics for Review

- n-armed bandit
- epsilon-greedy algorithms, softmax action selection
- MDPs, Bellman equations
- value iteration, policy iteration, do things asynchronously
- Monte Carlo policy evaluation

## 4. Vision

### 4.1. Chapter 1, 7, 11 (August 12)

#### Discussions

- Difference between perspective, weak perspective, orthographic (see table 1.1 on page 29)
- Decomposition of projection matrix into calibration and rotation/translation
- Essential and fundamental matrix derivations
- Binocular reconstruction – why only four constraints instead of six?
- Rectification – mainly makes finding correspondences easier, epipoles move to infinity
- Difference between correlation vs. edge matching – first one is correspondences between pixels, second one is higher level (between edges)
- Multi-scale edge matching – why second derivative filter? Because we want to match zero crossings (which are edges) to other zero crossings
- Structure from motion: isn't eq 8.2 good enough? (Hard optimization problem, optimizing other things can be easier)
- Uncalibrated weak-perspective - why does this work, since orthographic assumes points are on planes (it's because points are still in 3d, orthographic is just the projection method)

#### Topics for Review

- Perspective models
- Camera intrinsics and extrinsics
  - Intrinsics: center X/Y, scale X/Y, skew
  - Extrinsics: rotation, translation
- Camera calibration and recovery of intrinsics/extrinsics
- Essential/fundamental matrices
- Binocular reconstruction
- Local methods for binocular fusion (correlation between patches, multi-scale edge matching)
- Global methods for binocular fusion (essentially find energy function, minimize it)
- Structure from motion - set up constraints, solve non-linear least squares

- Weak calibration - estimate epipolar geometry, then do binocular stereopsis
- Eight point algorithm - compute F, decompose to find E, decompose to find R and t
- Uncalibrated weak-perspective - affine epipolar geometry/calibration
- Uncalibrated perspective cameras, bundle adjustment

### 4.2. Chapter 4, 9 (August 15)

#### Discussions

- Convolutions
- A high level view of aliasing, nyquist frequency and low pass filtering
- Image pyramids and use of scale representations
- Watershed transform and mean-shift

#### Topics for Review

- Table 4.1 (fourier transform chart)
- Aliasing
- Different clustering algorithms for segmentation

## 5. Machine Learning

### 5.1. Chapter 2-7, 17 (August 12)

#### Discussions

- Rundown of sum-product
- Junction tree

#### Topics for Review

- Independence and graphs
- Elimination - sum over all factors a variables appears in
- Reconstituted graph
- Moralization - resulting graph might have fewer independence assumptions
- Complexity of elimination is exponential in the size of the elimination clique
- Tree-structured elimination - sum-product/message passing
- Junction tree

### 5.2. Chapter 5-7, 10-12 (August 15)

#### Discussions

- Bayesians vs Frequentists overview
- LMS algorithm
- Generative models
- Iteratively reweighted least squares
- Brief overview of other classification methods
- EM : the incomplete likelihood vs expected conditional likelihood
- KL divergence perspective of EM

#### Topics for Review

- IRLS (and overview of other classification algorithms)
- EM
- Practice EM by derivations of conditional mixtures etc