Lecture 26: April 27, 2015
cs 573: Probabilistic Reasoning
Professor Nevatia
Spring 2015

Review

- Please complete course evaluation
- Exam2: April 29, class period, here (except for remote DEN students)
  - Closed book/notes
  - Detailed list of topics in following slide
  - Style similar to Exam1 but likely to be more descriptive
- Previous Lecture
  - Learning MN parameters
  - Intro to structure learning
- Today's objective
  - Review

Exam Topics

- Lec 13-25 (both inclusive); Exam1 topics will not be repeated in Exam 2
- Topics:
  - Gaussian Networks: Ch. 7, 14.2
  - Variational Approximation: 11.5 (only as covered in class)
  - MAP inference: 13.1, 13.2 and 13.3
  - Sampling: 12.1, 12.2 and 12.3 (including Metropolis-Hastings algorithm)
  - Temporal Models: 6.2, 15.1, 15.2, 15.3 (as covered in class)
  - Parameter Estimation: 17.1, 17.2 (exclude 17.2.4, 17.2.5), 17.3
  - Partially Observed Data: 19.1, 19.2 (except 19.2.2.5 and 19.2.2.6, 19.2.4)
  - Latent Dirichlet Allocation, class slides (Blei paper is optional)
  - Markov Networks: 20.1, 20.2, 20.3 (except 20.3.4)
  - Structure Learning18.1 to 18.4 (as covered in class)

Review of MN learning and Structure Learning

- Same slides as for Lec 25

Following Material is Just for Information: not included for Exam.

Making Decisions

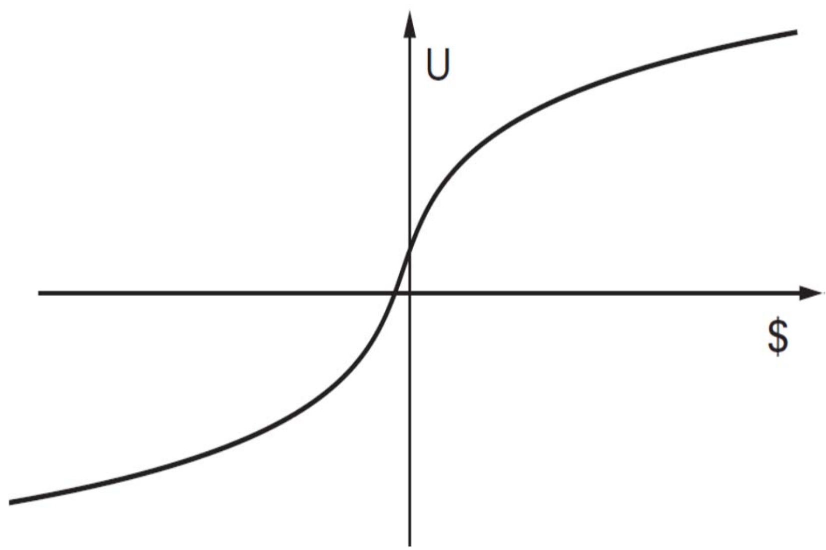
- Decision theory: what is the optimal action to take in a given situation?
- Compute probability distribution of the current state
- Evaluate likely results of possible actions and their *utility*
- Choose actions that maximize *expected* utility
- Computing utility may require looking ahead to see what the current state leads to.

An Example

- You have won \$1M in a game show, the host offers you a choice: either keep the \$1M or toss a coin, if heads come up you get \$3M otherwise you get nothing.
- Which choice would you make? Which is *rational*? Note that *Expected Monetary Value* of the coin toss choice is higher (\$1.5 M *vs* \$1M)
- An average person would choose to take \$1M. A very rich person may choose the coin toss
  - Decision reflects the utility function of the chooser
  - For an average person, utility of extra \$1M is high, value of additional \$2M is not so much higher

Utility of Money

- Monotonic: more is better, but not linear
  - An approximation is $U(s) = \log_2 (s + c)$, c is a constant
  - An example curve is shown in fig below (from human studies)
- *Risk averse* in the positive part of the curve (we accept a smaller sure reward, buy insurance…). Negative part is *risk seeking* (buy Lotto..)



MEU Principle

The principle of maximum expected utility (MEU principle) *asserts that, in a decision-making situation $\mathcal{D}$, we should choose the action a that maximizes the* expected utility:

$$\text{EU}[\mathcal{D}[a]] = \sum_{o \in \mathcal{O}} \pi_a(o) U(o).$$

- Game show example:
  - Accept: EU(Accept) = $0.5 \cdot U(S_k) + .5 \cdot U(S_{k+3M})$ (k- initial wealth) = $0.5 \cdot 0 + 0.5 \cdot \log_2(3M) \approx .5 \cdot 21.51 \approx 10.75$
  - Decline: EU(Decline) $= U(S_{k+1M}) = \log_2(1M) = 19.93$
  - EU of Accept is higher, if award exceeds $\$2^{40} \approx \$10^{32}$.
- Marginal utility of accept is even lower for a person with high wealth
- Preference for Accept implies a different utility function or an irrational decision

Course Review

- What have we studied?
- Graphical model representation (variables, relations amongst them)
  - Undirected and directed models
  - Template models
- Inference methods
  - Exact inference (Variable elimination, clique tree calibration)
  - Approximate inference (LBP, variational methods, sampling)
  - MAP inference
  - Temporal models (filtering, smoothing, maximum likelihood path)
  - Approximate methods for temporal models (particle filtering)

Course Review

- Learning
  - Maximum Likelihood Estimate
    - For BNs and MNs
      - With complete data
      - With incomplete data
  - Bayesian Parameter Estimation
  - Structure Learning

How can the techniques be used in practice?

- Have been applied to virtually all aspects of decision making
- Constructing an appropriate model remains important
  - What should be the variables in the node (especially, how to decide about hidden nodes that may be never observed)
  - Relations between nodes (links and parameters)
    - Directed or not
  - Discrete or continuous variables
  - …
- Separation between errors due to model construction and due to inference methods
  - More difficult when we have to use approximate inference methods
- Separation between inference and decisions