Inference with Graphical Models

Presented by - Amruta Purandare

Date - 09/14/2005

CS3710 Advanced AI

Overview

- Query Types
- Complexity of answering queries
- Simple Inferences for chain, loop n/ws
- Variable Elimination

Query I: Conditional Probability

- Given evidence: E=e
- Query: Y

- $P(Y=y|E=e) = P(Y=y, E=e) / P(E=e)$

- $P(Y=y, E=e) = \Sigma_w P(Y=y, E=e, w)$; $W=X-Y-E$

- $P(E=e) = \Sigma_y P(Y=y, E=e)$

Query II: Most Probable Explanation

- Given evidence: E=e
- Find: most likely assignment to ALL non-evidence variables; W=X-E

- $MPE(W|E=e) = \max_w P(w, E=e)$

- $MPE(A,W-A|E=e) \ != \max P(A|E=e)$

Query III: Maximum A Posteriori

- Given evidence: E=e
- Find: most likely assignment to only SOME non-evidence variables; Y=X-E-Z
- $MAP(Y|E=e) = \max_y \Sigma_z P(y,z|E=e)$
- Non-monotonic:
  MAP(Y1|E) != MAP(Y1,Y2|E=e)
- General Case of MPE

Summary

- Probability Query
  - Find: P(Y|E=e)
- MPE
  - Find: $\max_w P(w|E=e)$; W=X-E
- MAP
  - Find: $\max_y P(y|E=e)$; Y=X-E-Z

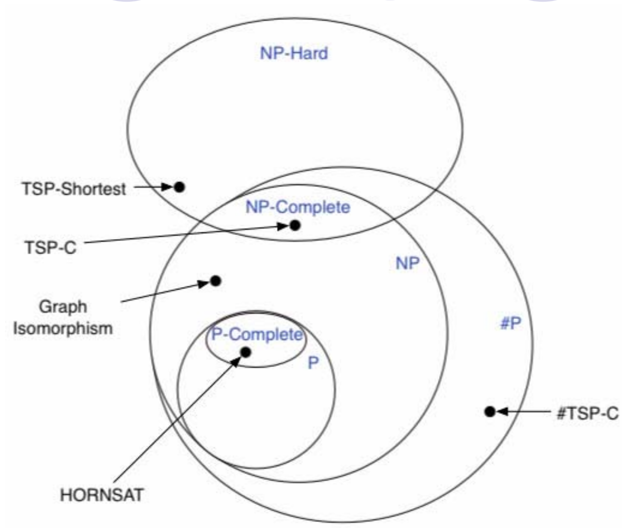
Complexity

Given BN β over X, where x ε val(X)

- Probability Query
  - Decide: $P(X=x)>0$ — NP-complete
  - Compute: $P(X=x)$ — #P-complete
- MPE:
  - If there exists x such that $P(x)>\delta$ — NP-complete
- MAP: — NP-hard

#P (Sharp-P)

- NP problems:
  - are there any solutions that satisfy given constraints?
    - e.g. 3SAT problem, subset-sum problem
- #P
  - How many solutions?
    - e.g. how many solutions satisfy a CNF formula
    - how many subsets of set S add to T



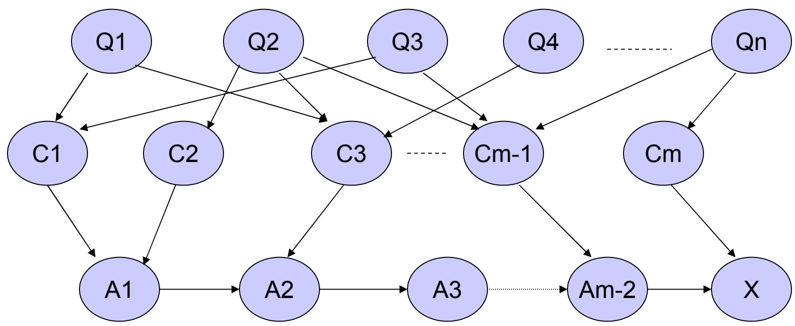
Background (from wikipedia)

- P: solvable in polynomial time
- P-complete: subclass of P, which all problems in P can be reduced to in polynomial time
- NP: solution can be guessed or verified in polynomial time
- NP-Hard: every NP is reducible to, in poly-time
- NP-complete: subclass of NP which is NP-Hard

  if you solve a NP-complete problem in poly-time, you can solve all NP problems in poly-time

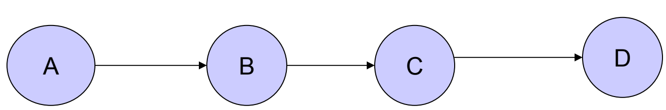
Theorem 6.2.1

- Decision problem $P(X=x)>0$ is NP-complete



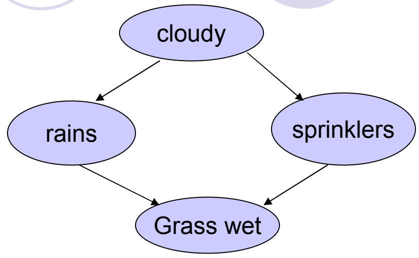
Simple Chain

- $P(B) = \Sigma_a P(a) P(B|a)$
- $P(C) = \Sigma_b P(b) P(C|b)$



- $P(X_{i+1}) = \Sigma_{xi} P(x_i) P(X_{i+1}|x_i)$ takes $O(nk^2)$
  k^2 multiplications, $k(k-1)$ additions
- $P(X_n) = \Sigma_{xn-1\ldots}\Sigma_{x2}\Sigma_{x1} P(x_1,x_2,\ldots x_{n-1},X_n)$ takes $O(k^n)$

Loop



- $p(W) = \Sigma_c \Sigma_s \Sigma_r \; p(W|r,s) \; p(r|c) \; p(s|c) \; p(c)$
- $= \Sigma_s \Sigma_r \; p(W|r,s) \; \Sigma_c \; p(r|c) \; p(s|c) \; p(c)$
- $= \Sigma_s \Sigma_r \; p(W|r,s) \; \mathcal{T}(r,s)$
- $\neq \Sigma_s \Sigma_r \; p(W|r,s) \; p(r) \; p(s)$

Factor Product

a1	b1	0.5
a1	b2	0.2
a2	b1	0.1
a2	b2	0.3
a3	b1	0.2
a3	b2	0.4

b1	c1	0.1
b1	c2	0.6
b2	c1	0.3
b2	c2	0.4

a1	b1	c1	0.5*0.1
a1	b1	c2	0.5*0.6
a1	b2	c1	0.2*0.3
a1	b2	c2	0.2*0.4
a2	b1	c1	0.1*0.1
a2	b1	c2	0.1*0.6
a2	b2	c1	0.3*0.3
a2	b2	c2	0.3*0.4
a3	b1	c1	0.2*0.1
a3	b1	c2	0.2*0.6
a3	b2	c1	0.4*0.3
a3	b2	c2	0.4*0.4

$ø(x,y,z) = ø1(x,y) \, . \, ø2(y,z)$

Factor Product

a1	b1	0.5
a1	b2	0.2
a2	b1	0.1
a2	b2	0.3
a3	b1	0.2
a3	b2	0.4

b1	c1	0.1
b1	c2	0.6
b2	c1	0.3
b2	c2	0.4

a1	b1	c1	0.5*0.1
a1	b1	c2	0.5*0.6
a1	b2	c1	0.2*0.3
a1	b2	c2	0.2*0.4
a2	b1	c1	0.1*0.1
a2	b1	c2	0.1*0.6
a2	b2	c1	0.3*0.3
a2	b2	c2	0.3*0.4
a3	b1	c1	0.2*0.1
a3	b1	c2	0.2*0.6
a3	b2	c1	0.4*0.3
a3	b2	c2	0.4*0.4

$\phi(x,y,z) = \phi1(x,y) \,.\, \phi2(y,z)$

Factor Product

a1	b1	0.5
a1	b2	0.2
a2	b1	0.1
a2	b2	0.3
a3	b1	0.2
a3	b2	0.4

b1	c1	0.1
b1	c2	0.6
b2	c1	0.3
b2	c2	0.4

a1	b1	c1	0.5*0.1
a1	b1	c2	0.5*0.6
a1	b2	c1	0.2*0.3
a1	b2	c2	0.2*0.4
a2	b1	c1	0.1*0.1
a2	b1	c2	0.1*0.6
a2	b2	c1	0.3*0.3
a2	b2	c2	0.3*0.4
a3	b1	c1	0.2*0.1
a3	b1	c2	0.2*0.6
a3	b2	c1	0.4*0.3
a3	b2	c2	0.4*0.4

$\phi(x,y,z) = \phi1(x,y) \,.\, \phi2(y,z)$

Factor Product

a1	b1	0.5
a1	b2	0.2
a2	b1	0.1
a2	b2	0.3
a3	b1	0.2
a3	b2	0.4

b1	c1	0.1
b1	c2	0.6
b2	c1	0.3
b2	c2	0.4

a1	b1	c1	0.5*0.1
a1	b1	c2	0.5*0.6
a1	b2	c1	0.2*0.3
a1	b2	c2	0.2*0.4
a2	b1	c1	0.1*0.1
a2	b1	c2	0.1*0.6
a2	b2	c1	0.3*0.3
a2	b2	c2	0.3*0.4
a3	b1	c1	0.2*0.1
a3	b1	c2	0.2*0.6
a3	b2	c1	0.4*0.3
a3	b2	c2	0.4*0.4

$\phi(x,y,z) = \phi 1(x,y) \, . \, \phi 2(y,z)$

Factor Marginalization

a1	b1	c1	0.2
a1	b1	c2	0.35
a1	b2	c1	0.4
a1	b2	c2	0.15
a2	b1	c1	0.5
a2	b1	c2	0.1
a2	b2	c1	0.3
a2	b2	c2	0.2
a3	b1	c1	0.25
a3	b1	c2	0.45
a3	b2	c1	0.15
a3	b2	c2	0.25

a1	c1	0.6
a1	c2	0.5
a2	c1	0.8
a2	c2	0.3
a3	c1	0.4
a3	c2	0.7

$\phi(x) = \Sigma_y \, \phi(x,y)$

Properties

- Commutative
  - $ø1.ø2 = ø2.ø1$
  - $\Sigma_x \Sigma_y ø = \Sigma_y \Sigma_x ø$
- Associative
  - $(ø1.ø2).ø3 = ø1.(ø2.ø3)$
- $\Sigma_x ø1ø2 = ø1 \Sigma_x ø2$ if $x \notin$ scope[ø1]

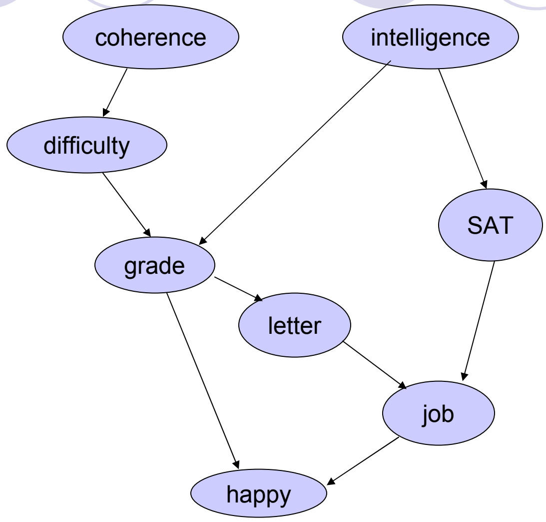
Variable Elimination

- $p(d) = \Sigma_c \Sigma_b \Sigma_a p(a,b,c,d)$
- $= \Sigma_c \Sigma_b \Sigma_a ø_a ø_b ø_c ø_d$
- …assuming

  scope[$ø_a$]={a}, scope[$ø_b$]={a,b}, scope[$ø_c$]={b,c}

  scope[$ø_d$]={c,d}
- $= \Sigma_c \Sigma_b ø_c ø_d \Sigma_a ø_a ø_b$
- $= \Sigma_c ø_d \Sigma_b ø_c \Sigma_a ø_a ø_b$

Sum-Product



Cont…

p(c,d,i,g,l,s,j,h)

$= p(c)\ p(i)\ p(d|c)\ p(g|i,d)\ p(s|i)\ p(l|g)\ p(j|l,s)\ p(h|g,j)$

$= ø(c)\ ø(i)\ ø(d,c)\ ø(g,i,d)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)$

Cont…

$p(J) = \Sigma_l\Sigma_s\Sigma_g\Sigma_h\Sigma_i\Sigma_d\Sigma_c\ ø(c)\ ø(i)\ ø(d,c)\ ø(g,i,d)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)$

Eliminating c,

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\Sigma_i\Sigma_d ø(i)\ ø(g,i,d)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \Sigma_c\ ø(c)ø(d,c)$

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\Sigma_i\Sigma_d ø(i)\ ø(g,i,d)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \tau(d)$

Cont…

Eliminating d,

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\Sigma_i\ ø(i)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \Sigma_d\ \tau(d)\ ø(g,i,d)$

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\Sigma_i\ ø(i)\ ø(s,i)\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \tau(g,i)$

Eliminating i,

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \Sigma_i\ ø(i)\ ø(s,i)\ \tau(g,i)$

$= \Sigma_l\Sigma_s\Sigma_g\Sigma_h\ ø(l,g)\ ø(j,l,s)\ ø(h,g,j)\ \tau(s,g)$

Cont…

Eliminating h,

= $\Sigma_l \Sigma_s \Sigma_g$ ø(l,g) ø(j,l,s) τ(s,g) Σ_h ø(h,g,j)

= $\Sigma_l \Sigma_s \Sigma_g$ ø(l,g) ø(j,l,s) τ(s,g) τ(g,j)

Eliminating g,

= $\Sigma_l \Sigma_s$ ø(j,l,s) Σ_g ø(l,g) τ(s,g) τ(g,j)

= $\Sigma_l \Sigma_s$ ø(j,l,s) τ(l,s,j)

Eliminating s,

= Σ_l τ(j,l)

Eliminating l,

= τ(j)

Restricted Factors - when given evidence

- Used for calculating conditional prob. queries

$$P(Y'|U=u) = \frac{P(Y',U=u)}{P(U=u)} = \frac{P(Y',U=u)}{\sum_y P(Y'=y,U=u)}$$

- Notation: $ø_{|U=u}(Y')= ø(Y',u)$ where Y'=Y-U
- Student Example revisited:
  - P(C,D,I,G,S,L,H)

    = ø(C) ø(I) ø(D,C) ø(G,I,D) ø(S,I) ø(L,G) ø(J,L,S) ø(H,G,J)
  - Given I=i, H=h,

    P(C,D,I=i,G,S,L,J,H=h)

    = ø(C) $ø_{|I=i}$() ø(D,C) $ø_{|I=i}$(G,D) $ø_{|I=i}$(S) ø(L,G) ø(J,L,S) $ø_{|H=h}$(G,J)