Inference of SIG
-- Inside Outside Algorithm –

Song-Chun Zhu, Sinisa Todorovic, and Ales Leonardis

At CVPR, Providence, Rhode Island
June 16, 2012

Inside-Outside Algorithm

- Notation:
  - Let $\mathbf{O} = O_1, O_2, ..., O_T$ be the observation sequence generated by a SCFG *G*.
  - let i, j, k be integer numbers corresponding to each of the non-terminal symbols.
  - Let m be an integer corresponding to a terminal symbol.
  - The grammar *G* has the Chomsky Normal Form

    $i \rightarrow jk$ $\qquad$ $i \rightarrow m$

Inside-Outside Algorithm

- Notation:
  - Probability matrix **A** and **B**

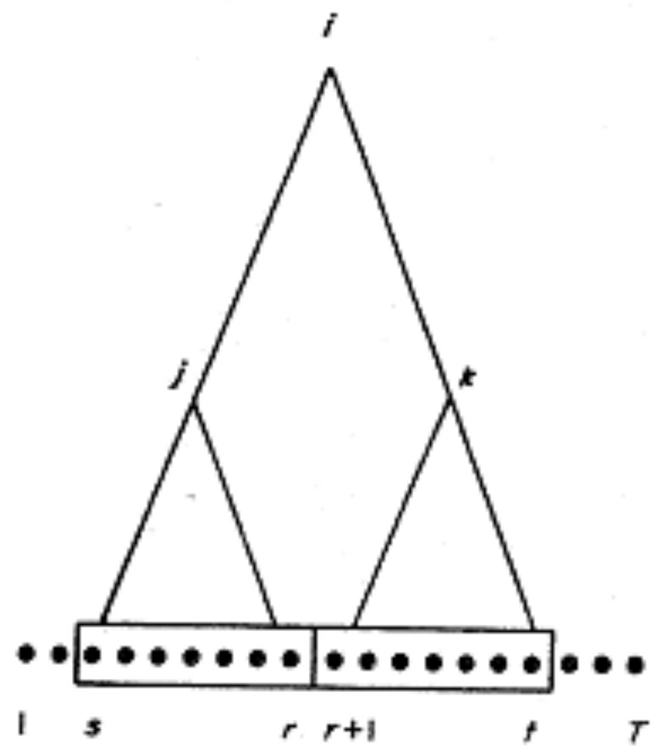
    $$a[i, j, k] = P(i \to jk)$$

    $$b[i, m] = P(i \to m)$$

  - $a[i, j, k]$ is the probability that the non-terminal symbol i generate the pair of non-terminal symbols j and k
  - $b[i, m]$ is the probability that the non-terminal symbol i generate a single terminal symbol m

Inner Probability

$e(s,t,i)$ = Probability of the non-terminal symbol i generating the observation $O(s),...,O(t)$



calculation of inner probabilities

Computation of Inner Probability

- When $s = t$

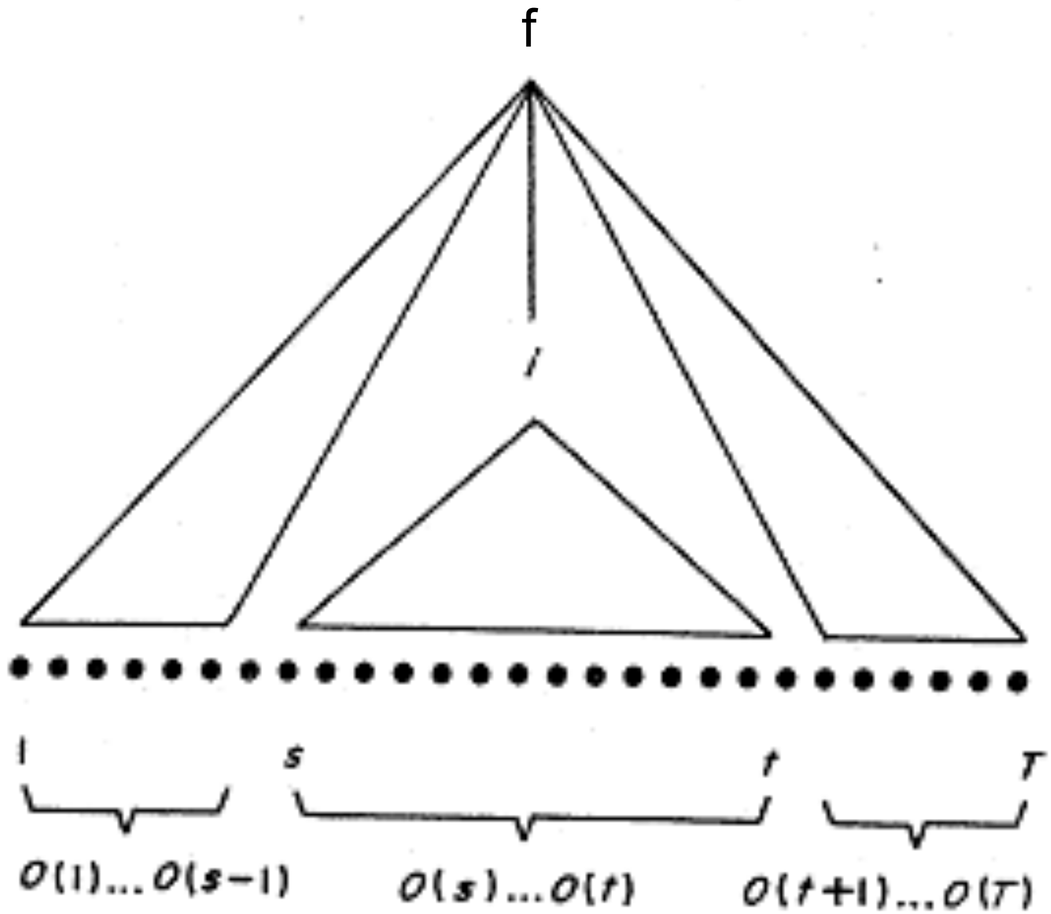
$$e(s,s,i) = P(i \rightarrow O(s)) = b[i,O(s)]$$

- When $s \neq t$

$$e(s,t,i) = \sum_{j,k} \sum_{r=s}^{t-1} a[i,j,k] e(s,r,j) e(r+1,t,k)$$

- The quantity *e* can be computed recursively by determining *e* for all sequences of length 1, then all sequences of length 2, and so on.

Outer Probability



Outer Probability

- Define outer probability f as

$$f(s,t,i) = P(S \Rightarrow O(1)...O(s-1), i, O(t+1)...O(T))$$

- f can be computed by

$$f(s,t,i) = \sum_{j,k} [\sum_{r=1}^{s-1} f(r,t,j) a[j,k,i] e(r,s-1,k)$$

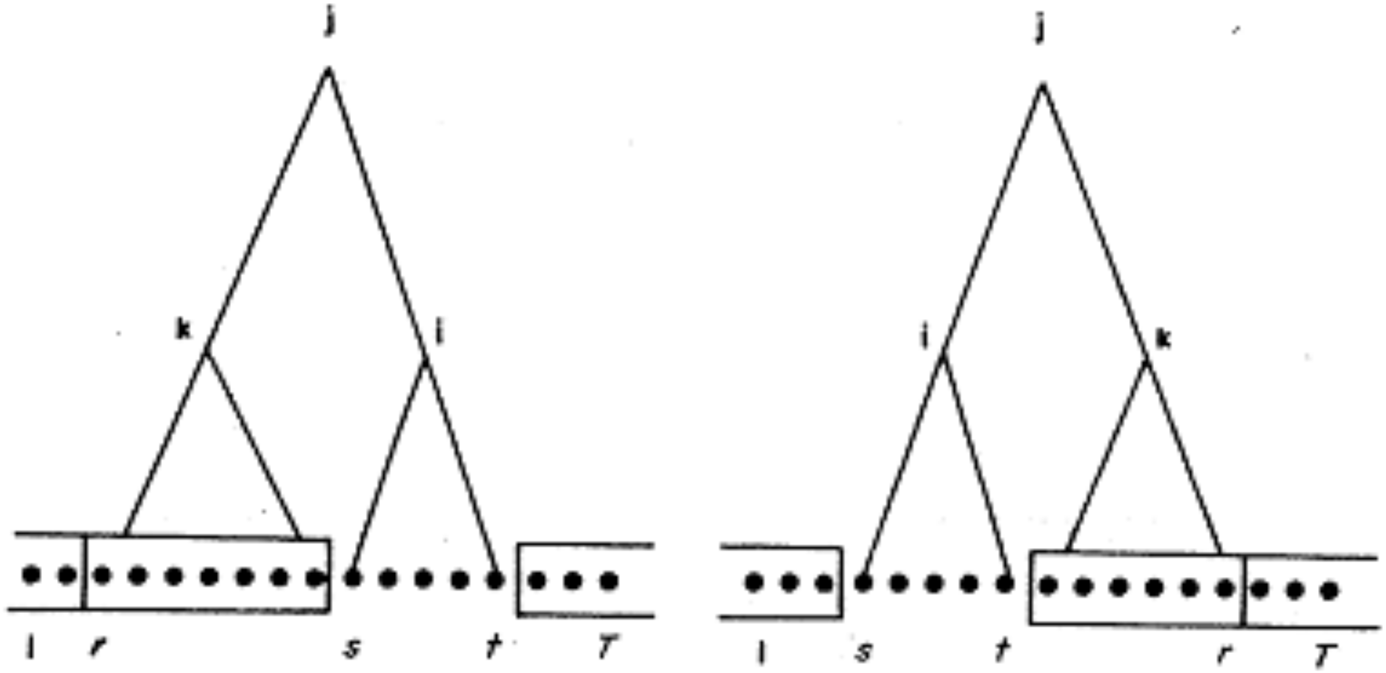
$$+ \sum_{r=t+1}^{T} f(s,r,j) a[j,i,k] e(t+1,r,k)]$$

- And

$$f(1,T,i) = \begin{cases} 1 & \text{if } (i = S) \\ 0 & \text{otherwise} \end{cases}$$

Computation of Outer Probability

Because a non-terminal node allows binary splits
$i \rightarrow j\ k$

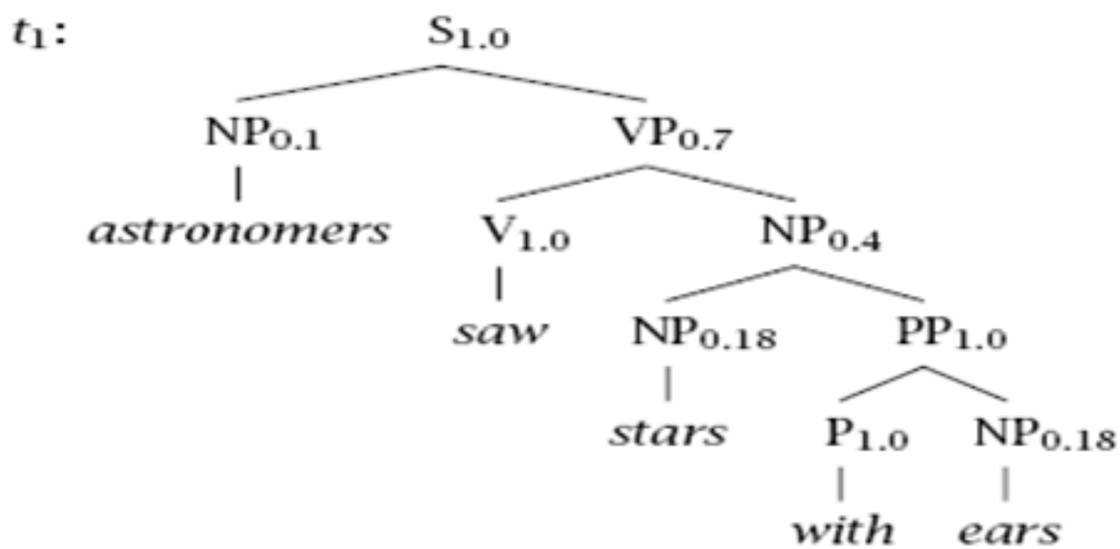


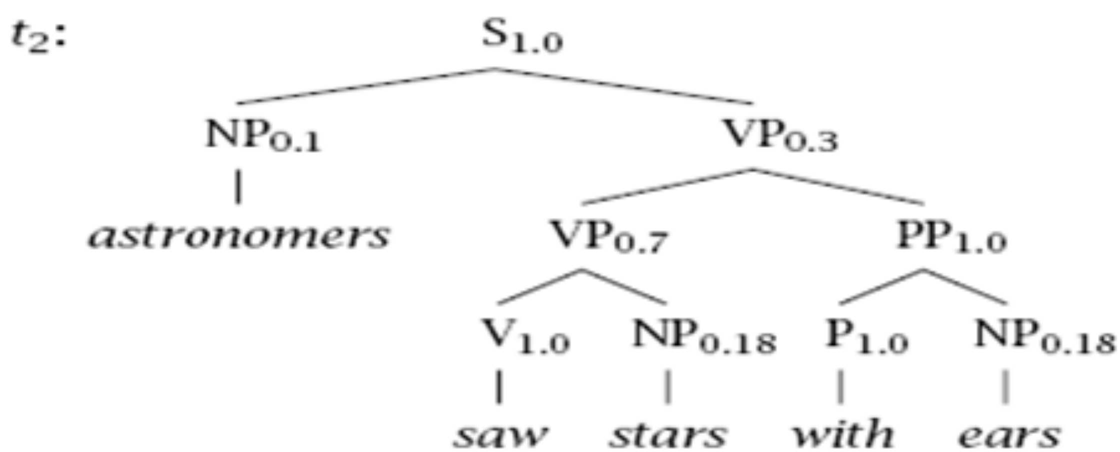
Inside-Outside Algorithm: Example

$S \rightarrow NP\ VP$	1.0	$NP \rightarrow NP\ PP$	0.4
$PP \rightarrow P\ NP$	1.0	$NP \rightarrow$ *astronomers*	0.1
$VP \rightarrow V\ NP$	0.7	$NP \rightarrow$ *ears*	0.18
$VP \rightarrow VP\ PP$	0.3	$NP \rightarrow$ *saw*	0.04
$P \rightarrow$ *with*	1.0	$NP \rightarrow$ *stars*	0.18
$V \rightarrow$ *saw*	1.0	$NP \rightarrow$ *telescopes*	0.1

x = astronomers saw stars with ears

Inside-Outside Algorithm: Example





Inside-Outside Algorithm: Example

- e(1,1,S)=0 e(1,1,PP)=0 e(1,1,VP)=0 e(1,1,NP)=0.1 e(1,1,V)=0 e(1,1,P)=0
- e(2,2,S)=0 e(2,2,PP)=0 e(2,2,VP)=0 e(2,2,NP)=0 .04 e(2,2,V)=1 e(2,2,P)=0
- e(3,3,S)=0 e(3,3,PP)=0 e(3,3,VP)=0 e(3,3,NP)=0.18 e(3,3,V)=0 e(3,3,P)=0
- e(4,4,S)=0 e(4,4,PP)=0 e(4,4,VP)=0 e(4,4,NP)=0 e(4,4,V)=0 e(4,4,P)=1
- e(5,5,S)=0 e(5,5,PP)=0 e(5,5,VP)=0 e(5,5,NP)=0.18 e(5,5,V)=0 e(5,5,P)=0

- e(1,2,S)=0 e(1,2,PP)=0 e(1,2,VP)=0 e(1,2,NP)=0 e(1,2,V)=0 e(1,2,P)=0
- e(2,3,S)=0 e(2,3,PP)=0 e(2,3,VP)=0 e(2,3,NP)=0 e(2,3,V)=0 e(2,3,P)=0
- ……
- f(1,5,S)=1 f(2,5,VP)=f(1,5,S)*a[S,NP,VP]*e(1,1,NP)=0.1
- f(3,5,NP)=f(2,5,VP))*a[VP,V,NP]*e(2,2,V)=0.07

- P(O|G) = 0.015876