CSE 322:
Introduction to Formal
Models in Computer Science

Cocke-Kasami-Younger Algorithm
Paul Beame

Determining whether \( w \in L(G) \)

- Assume \( G=(V, \Sigma, R, S) \) is in Chomsky Normal Form
- Grammar rules allowed:
  - \( A \to BC \) where \( B, C \in V \), \( B, C \neq S \)
  - \( A \to a \) where \( a \in \Sigma \)
  - \( S \to \epsilon \)
- If \( w=\epsilon \) check whether \( S \to \epsilon \) is in \( R \)
- If \( w=a \in \Sigma \) then check whether \( S \to a \) is in \( R \)
- Otherwise, parse tree must be a binary tree and first rule is some \( S \to BC \)

Recursive Algorithm (Exponential Time)

Generates(A,w)
if \( |w| \leq 1 \) output true iff \( A \to w \) is a rule in \( R \)
else
  \( n \leftarrow |w| \)
  for \( k=1 \) to \( n-1 \)
      \( x \leftarrow w[1..k] \); \( y \leftarrow w[k+1..n] \)
      for each rule \( A \to BC \) in \( R \)
          if Generates(B,x) and Generates(C,y)
              output true
      endfor
  endfor
output false
endif

Dynamic Programming

- All the recursive calls are subproblems of the type Generates(A,x) where
  - \( A \in V \)
  - \( x=w[i..j] \)
  - Intervals in \( w \) get shorter the deeper the call
- CKY Algorithm: Create a table whose \((i,j)\)th entry is the list of all variables that can generate the string \( w[i..j] \)
  - Fill out table starting with short intervals first
  - Answer is whether \( S \) is in table(1,n) where \( n=|w| \)

CKY algorithm: \( O(n^3) \) time

- Base
  - for all \( i=1 \) to \( n \)
  - table(i,j)←\{variables \( A \) with rule \( A \to w \)\}
- Iteration for \( d=1 \) to \( n-1 \)
  - Entries table(i,j) with \( j-i<d \) already computed
  - for every \((i,j)\) with \( j-i+d \) do
    - for \( k=1 \) to \( j-1 \)
      - for every rule \( A \to BC \)
        - if \( B=table(i,k) \) and \( C=table(k+1,j) \)
          - Add \( A \) to table(i,j)