

# 合一算法

# Unification

詹卫东

[zwd@pku.edu.cn](mailto:zwd@pku.edu.cn)

$$F = \begin{pmatrix} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{pmatrix}$$

特征结构的扩展表示



$$F = \begin{pmatrix} \text{CONTENT} & \begin{pmatrix} \text{NUMBER} & \begin{pmatrix} \text{CONTENTS} & \text{SG} \\ \text{POINTER} & \text{NULL} \end{pmatrix} \\ \text{PERSON} & \begin{pmatrix} \text{CONTENTS} & 3 \\ \text{POINTER} & \text{NULL} \end{pmatrix} \end{pmatrix} \\ \text{POINTER} & \text{NULL} \end{pmatrix}$$

# 特征结构的数据表示

为在合一运算中统一表达不同的特征结构，原特征结构进行如下改造：

(1) 每个特征结构由两个字段content和pointer构成。

(2) 对于原特征结构F的一个特征项和特征值对儿 [A:V]，

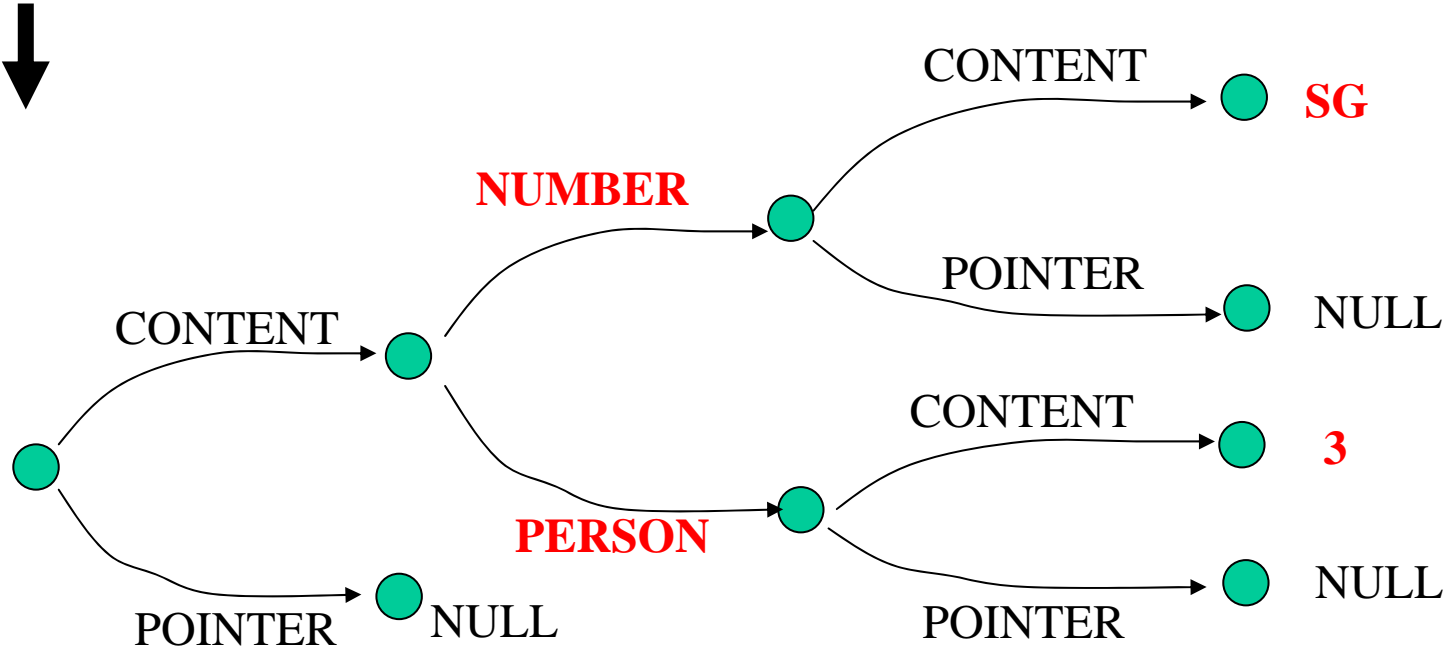
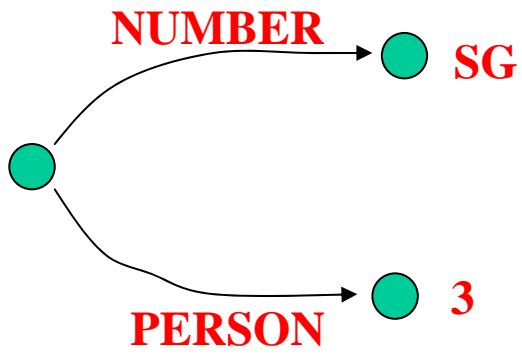
如果V为原子或另一个特征结构，则新的特征结构 (NF) 为：

$$\left[ \begin{array}{l} \text{content : A} \left[ \begin{array}{l} \text{content : V} \\ \text{pointer: null} \end{array} \right] \\ \text{pointer: null} \end{array} \right]$$

如果V为一个跟其他特征共享值的特征结构，则新的特征结构 (NF) 为：

$$\left[ \begin{array}{l} \text{content: A} \left[ \begin{array}{l} \text{content : null} \\ \text{pointer: [V-pointer]} \end{array} \right] \\ \text{pointer: null} \end{array} \right]$$

V-pointer指向v的共享值，通常用① ② .....等表示。

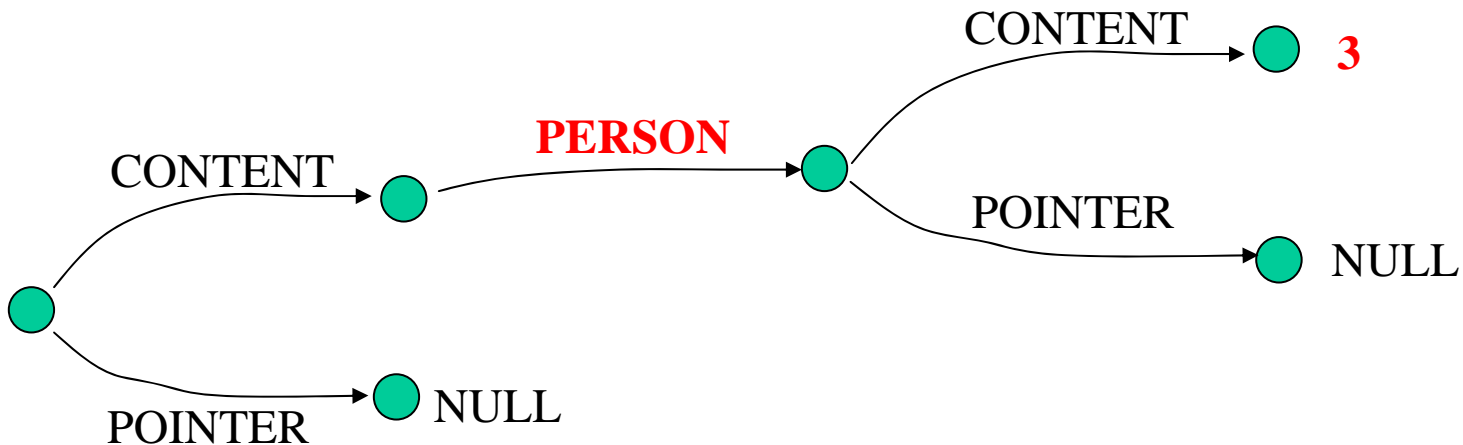
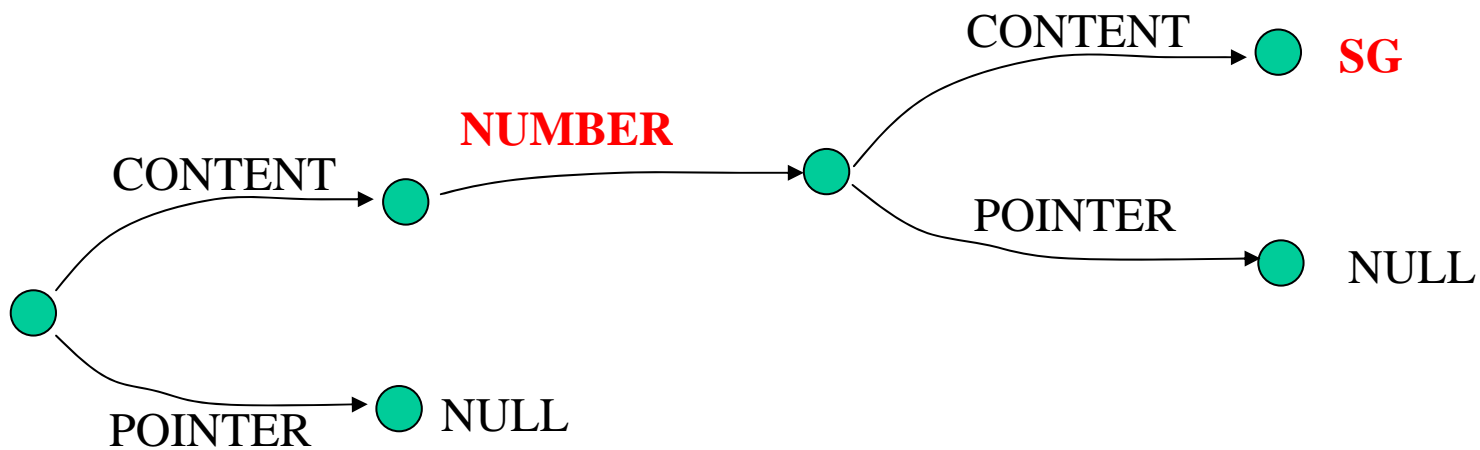


# 合一算法

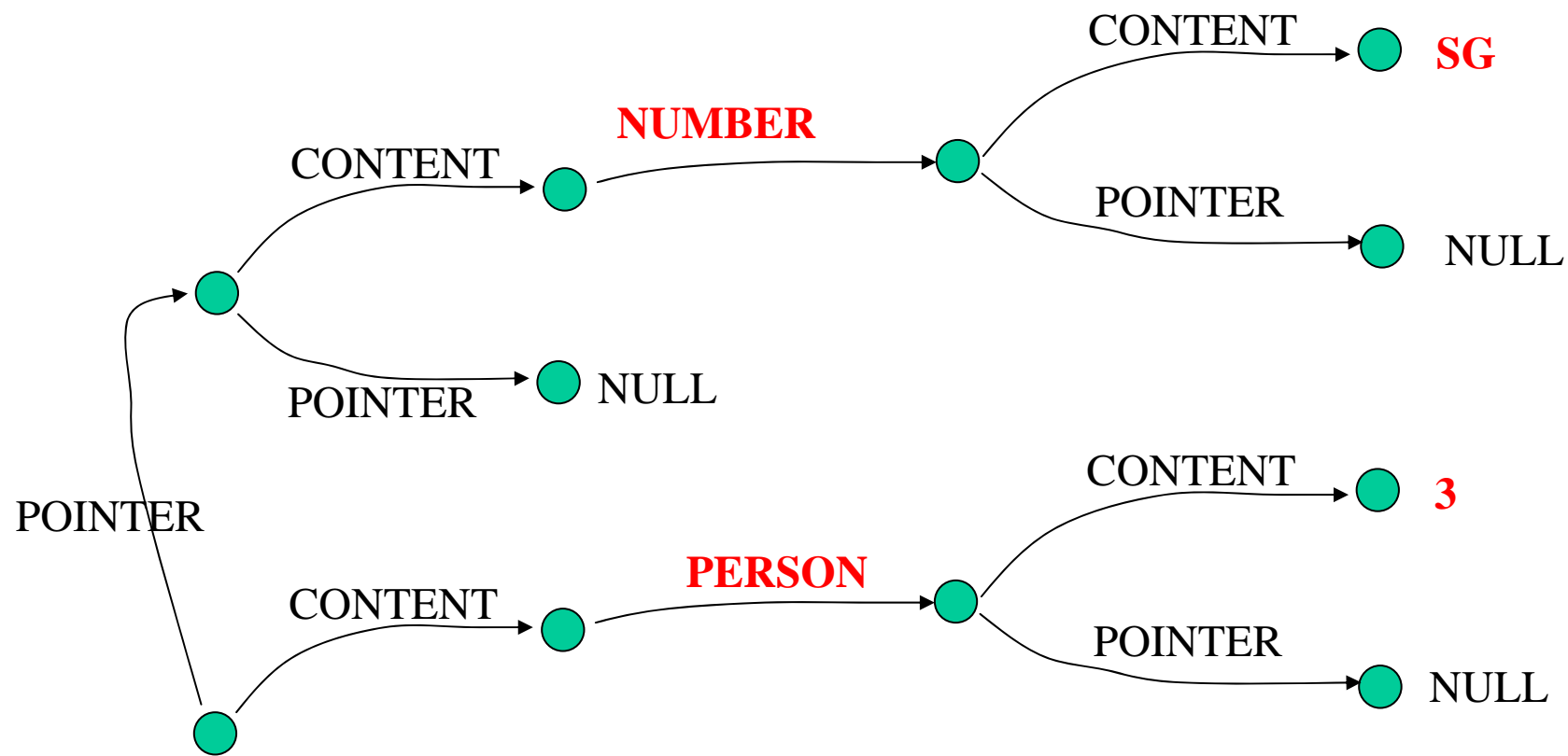
function **UNIFY**(f1, f2)

1. f1-real  $\leftarrow$  f1.content
2. f2-real  $\leftarrow$  f2.content
3. **if** f1-real is null **then**
4.     f1.pointer  $\leftarrow$  f2
5.     **return** f2
6. **else if** f2-real is null **then**
7.     f2.pointer  $\leftarrow$  f1
8.     **return** f1
9. **else if** f1-real and f2-real are identical **then**
10.     f1.pointer  $\leftarrow$  f2
11.     **return** f2
12. **else if** both f1-real and f2-real are complex feature structures **then**
13.     f2.pointer  $\leftarrow$  f1
14.     **for** each feature in f2-real **do**
15.         f1-feature  $\leftarrow$  Find or create a feature corresponding to feature in f1-real
16.         **if** **UNIFY**(f1.feature.value, f2.feature.value) == failure **then**
17.             **return** failure
18.     **return** f1
19. **else return** failure

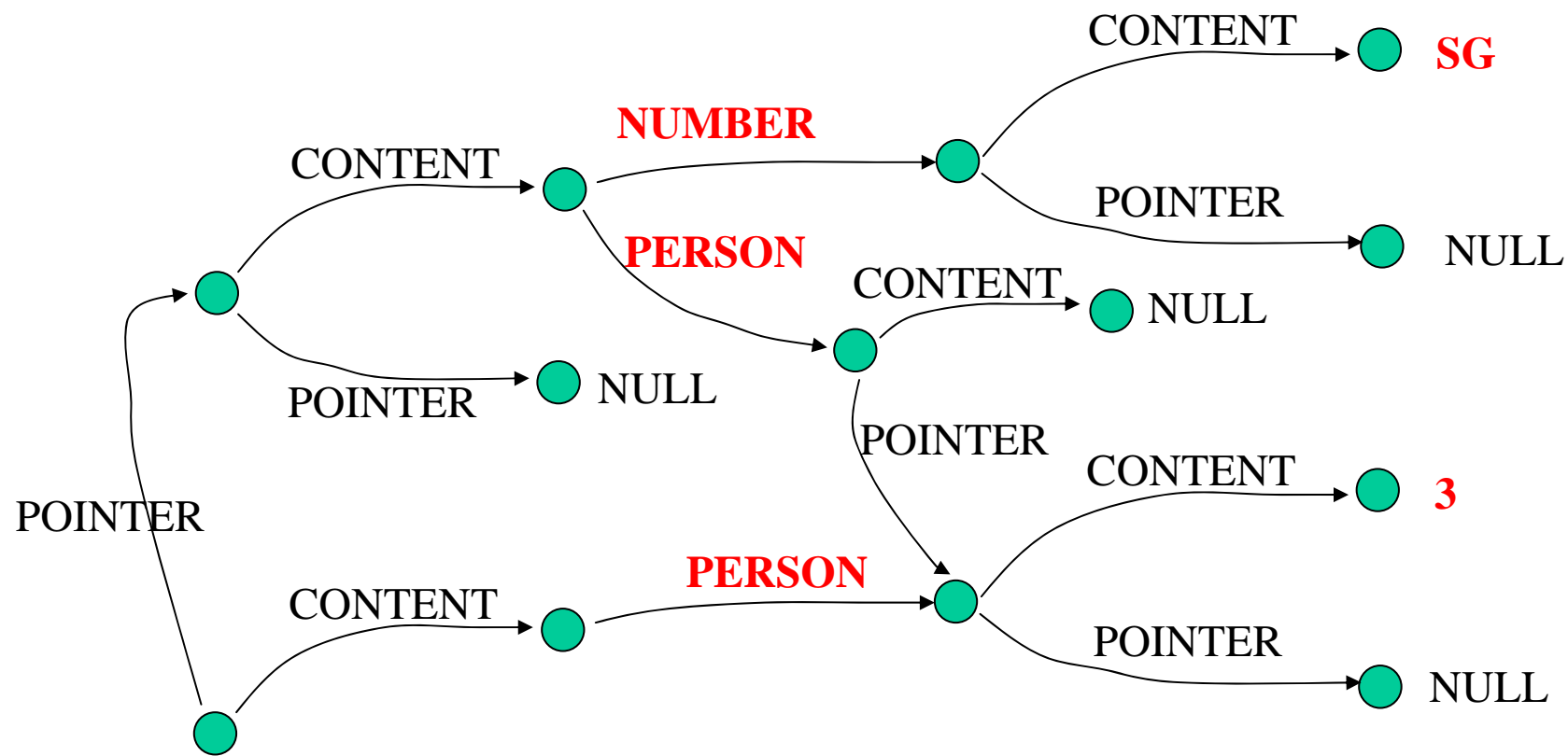
$$\left( \begin{array}{cc} \text{NUMBER} & \text{SG} \end{array} \right) \cup \left( \begin{array}{cc} \text{PERSON} & 3 \end{array} \right) = \left( \begin{array}{cc} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{array} \right)$$



$$\left( \begin{array}{cc} \text{NUMBER} & \text{SG} \end{array} \right) \cup \left( \begin{array}{cc} \text{PERSON} & 3 \end{array} \right) = \left( \begin{array}{cc} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{array} \right)$$



$$\left( \begin{array}{cc} \text{NUMBER} & \text{SG} \end{array} \right) \cup \left( \begin{array}{cc} \text{PERSON} & 3 \end{array} \right) = \left( \begin{array}{cc} \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{array} \right)$$



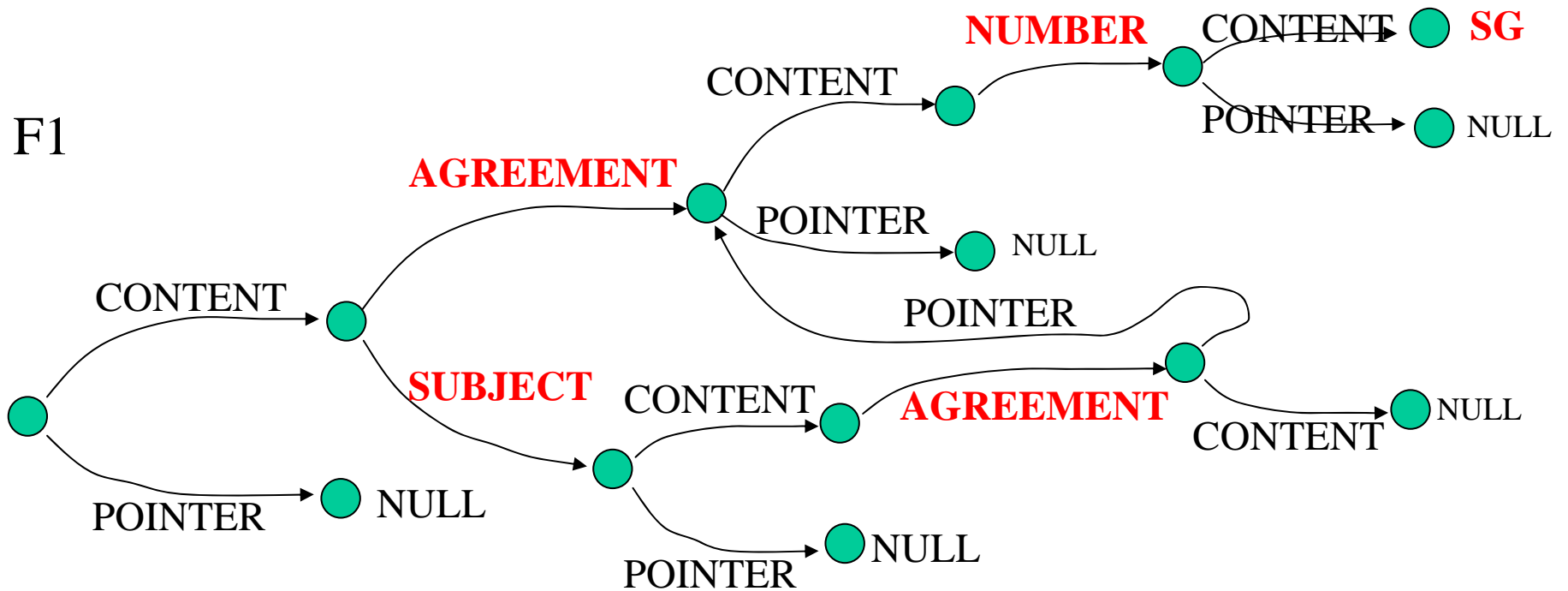
$$F1 = \left[ \begin{array}{l} \text{AGREEMENT} \quad \boxed{1} \left[ \begin{array}{l} \text{NUMBER} \quad \text{SG} \end{array} \right] \\ \text{SUBJECT} \quad \left[ \begin{array}{l} \text{AGREEMENT} \quad \boxed{1} \end{array} \right] \end{array} \right]$$

U

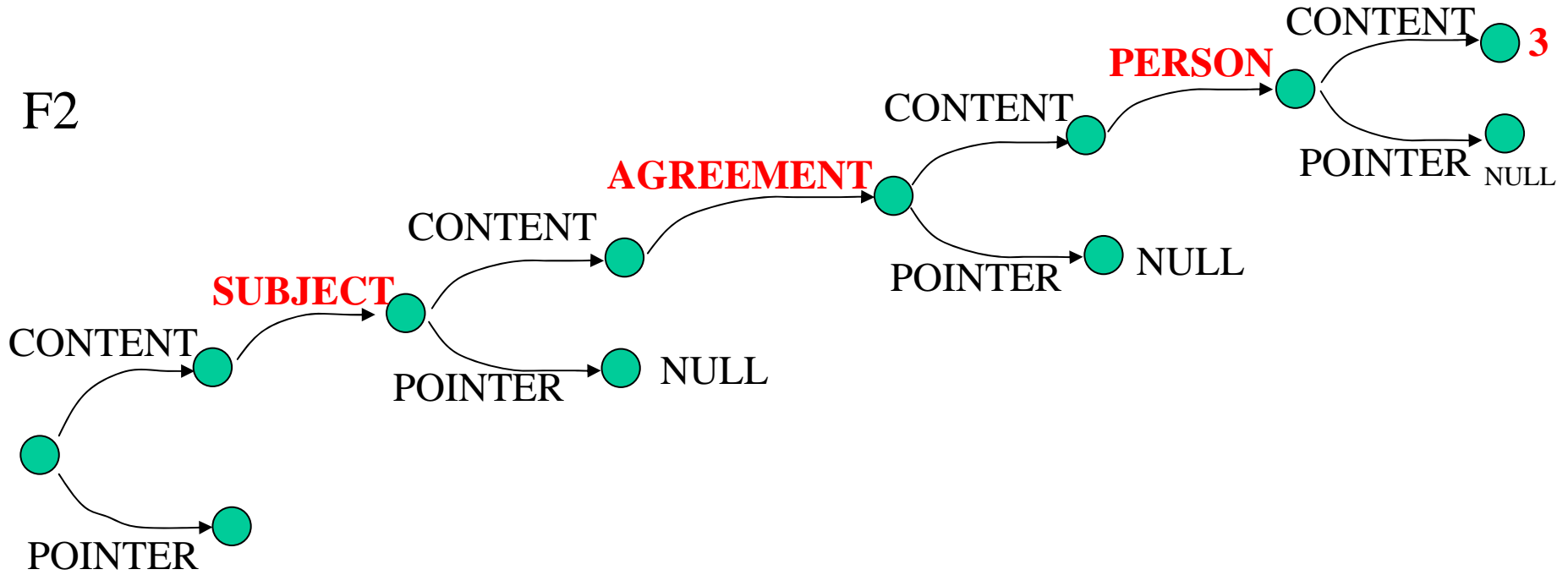
$$F2 = \left[ \begin{array}{l} \text{SUBJECT} \quad \left[ \begin{array}{l} \text{AGREEMENT} \quad \left[ \begin{array}{l} \text{PERSON} \quad 3 \end{array} \right] \end{array} \right] \end{array} \right]$$

$$= \left[ \begin{array}{l} \text{AGREEMENT} \quad \boxed{1} \left[ \begin{array}{l} \text{NUMBER} \quad \text{SG} \\ \text{PERSON} \quad 3 \end{array} \right] \\ \text{SUBJECT} \quad \left[ \begin{array}{l} \text{AGREEMENT} \quad \boxed{1} \end{array} \right] \end{array} \right]$$

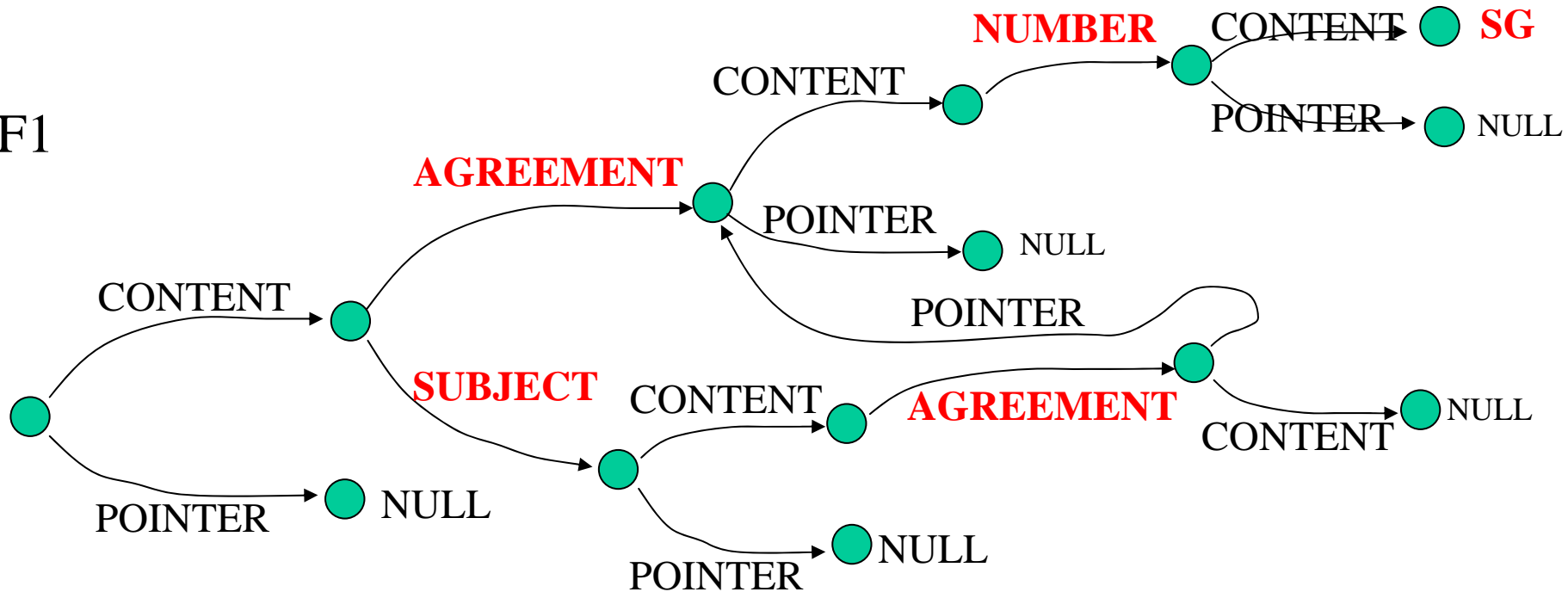
$$F1 = \left[ \begin{array}{cc} \text{AGREEMENT} & \boxed{1} \\ \text{SUBJECT} & \end{array} \right] \left[ \begin{array}{cc} \text{NUMBER} & \text{SG} \\ \text{AGREEMENT} & \boxed{1} \end{array} \right]$$



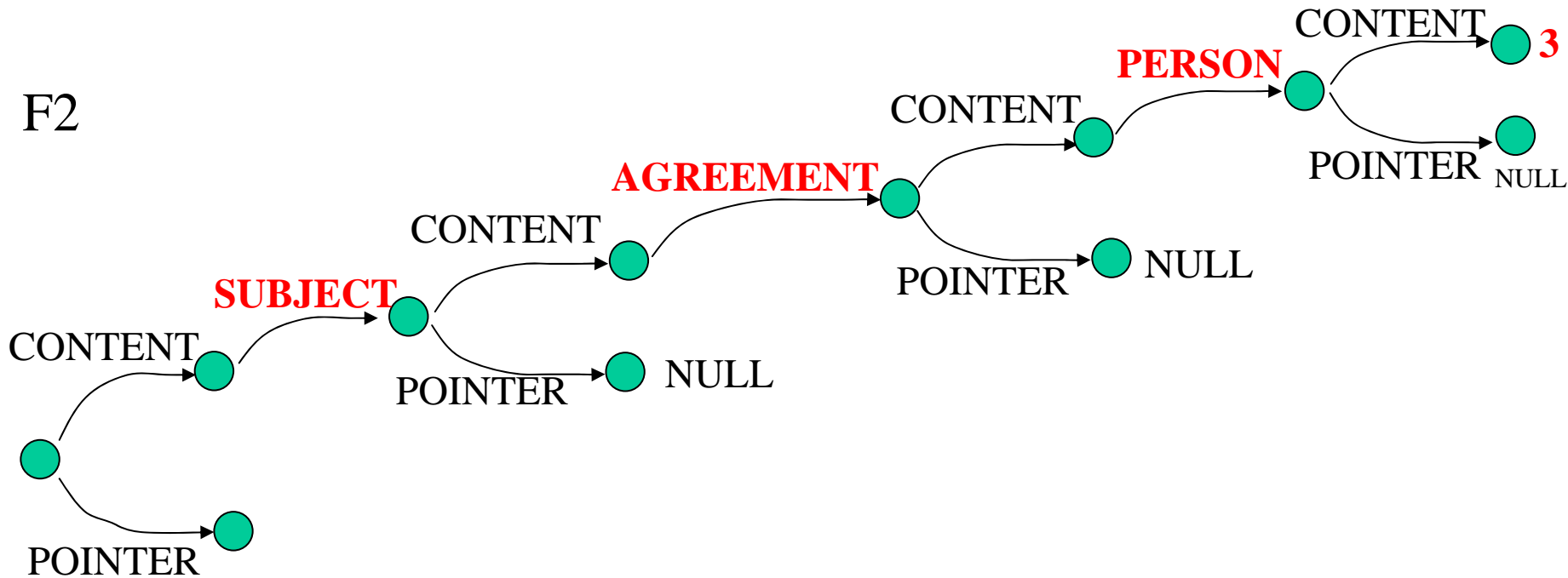
F2= [ SUBJECT [ AGREEMENT [ PERSON 3 ] ] ]



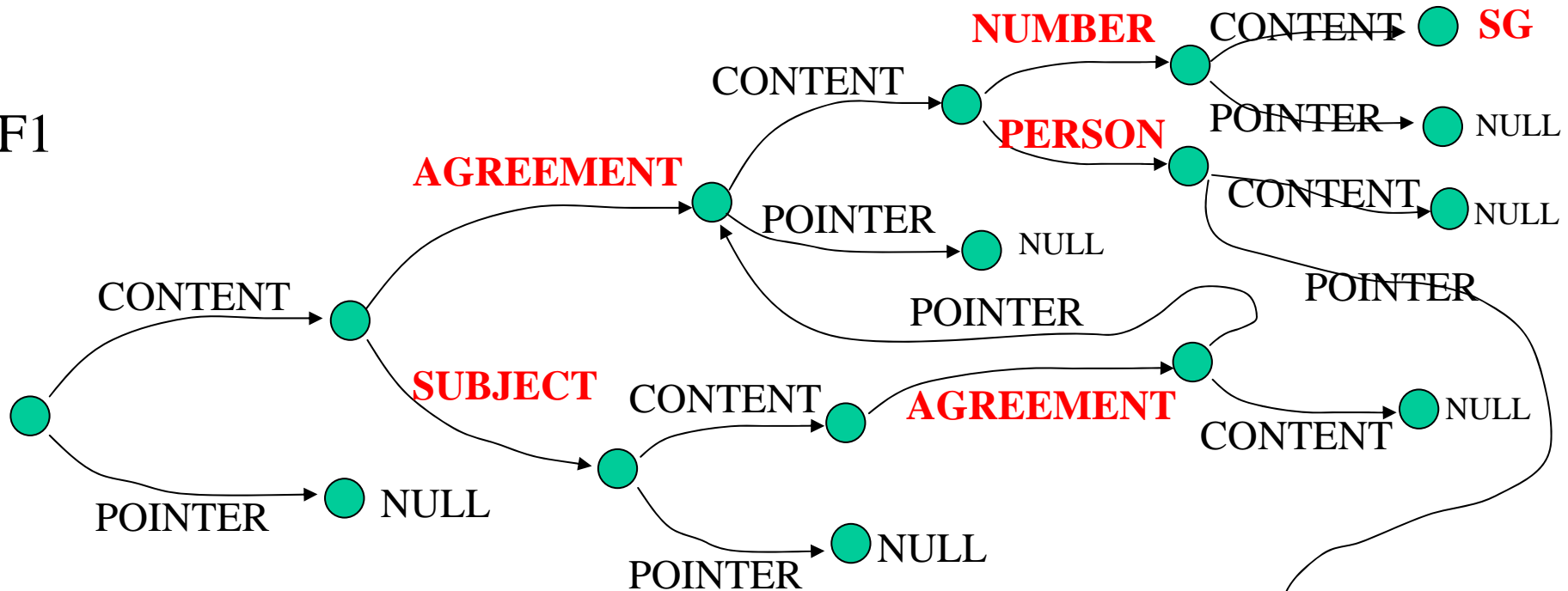
F1



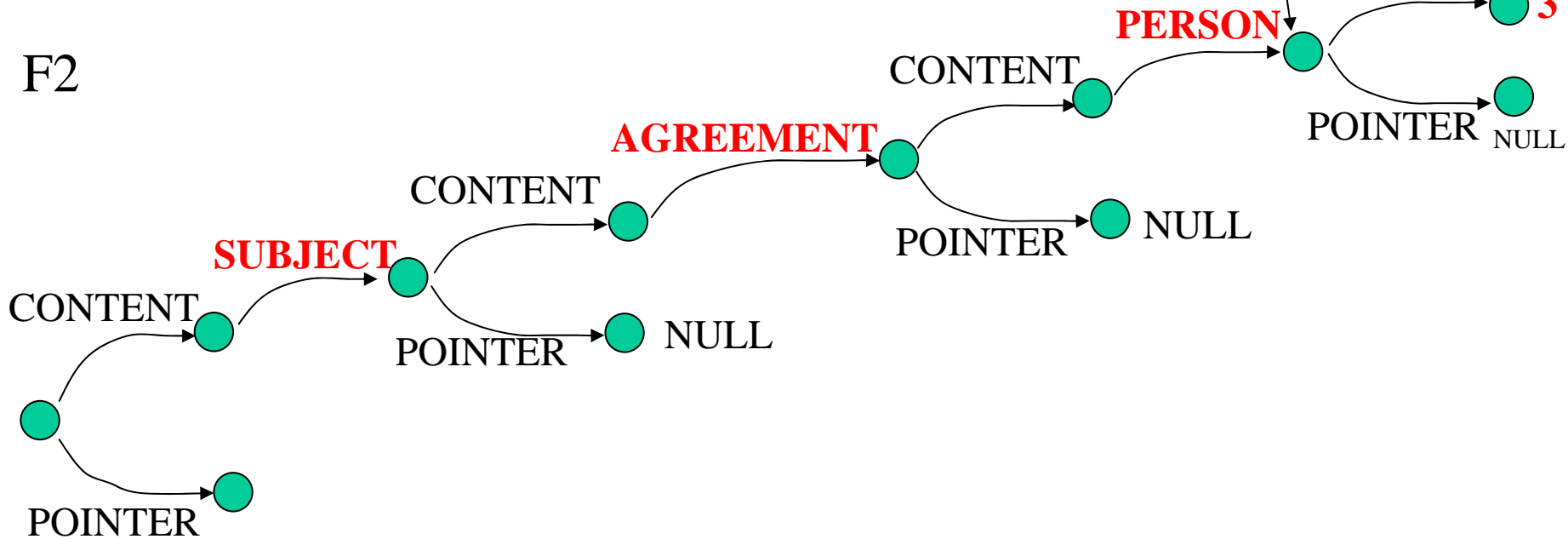
F2

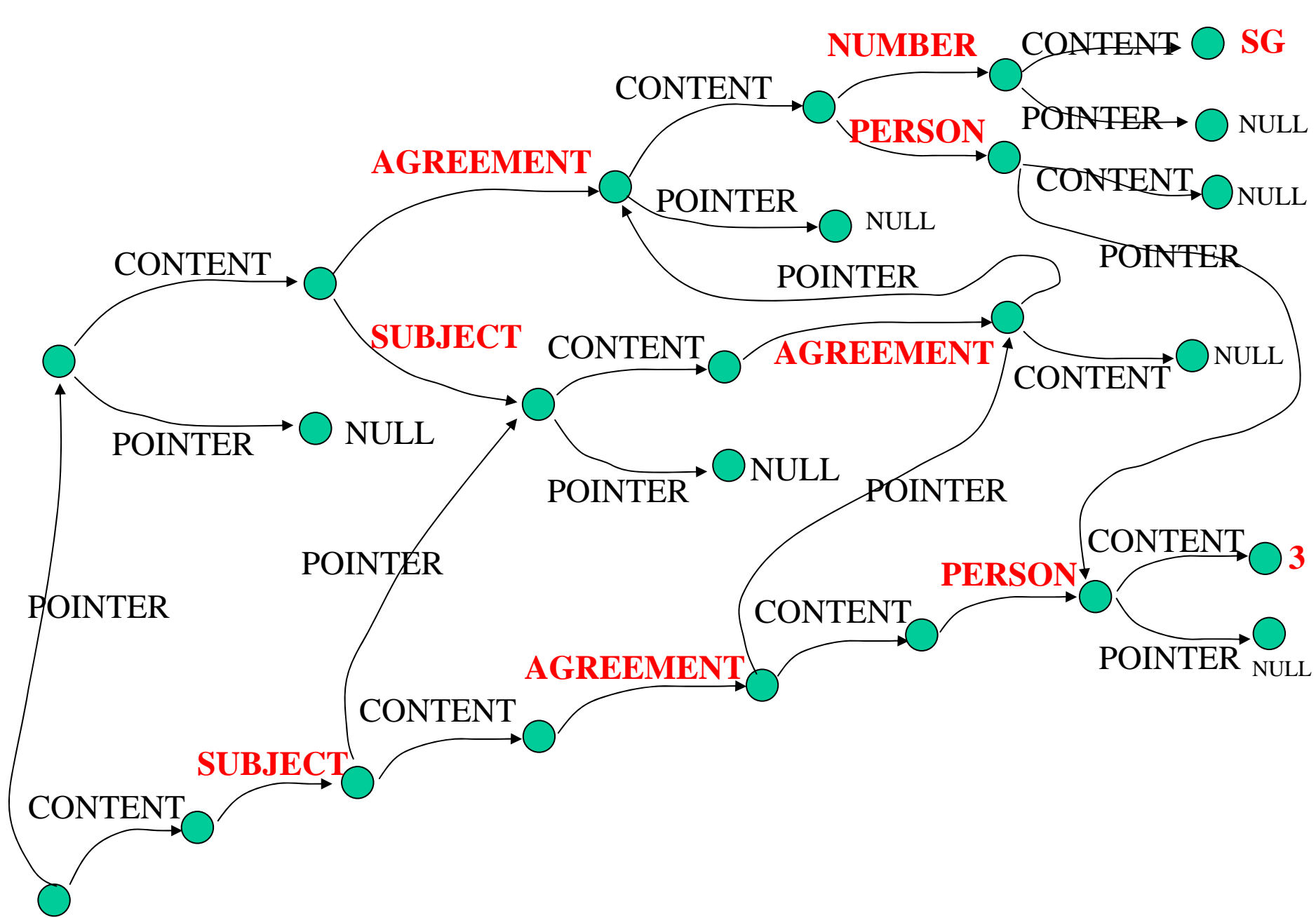


F1



F2





# 参考文献

- 冯志伟、孙乐译 《自然语言处理综论》  
电子工业出版社2005年版。第11章。  
Daniel Jurafsky & James H. Martin, 2000,  
Speech and Language Processing, Pearson  
Education, Inc., Prentice Hall.