

LÝ THUYẾT TÍNH TOÁN

BÀI 6: VĂN PHẠM PHI NGỮ CẢNH

Phạm Xuân Cường
Khoa Công nghệ thông tin
cuongpx@tlu.edu.vn

1. Khái niệm
2. Định nghĩa hình thức
3. Văn phạm nhập nhằng
4. Dạng chuẩn tắc Chomsky

Khái niệm

- Văn phạm phi ngữ cảnh = Context-free Grammar (CFG)
- **CFG**: Là một phương pháp mạnh hơn để mô tả ngôn ngữ
- Ứng dụng:
 - Bộ biên dịch trong các ngôn ngữ lập trình
 - Bộ phân tích trong các trình biên dịch và thông dịch
- Ví dụ:
$$E \rightarrow E + T \mid T$$
$$T \rightarrow T \times F \mid F$$
$$F \rightarrow (E) \mid a$$

Một văn phạm gồm có:

- Tập các **quy tắc thay thế** \equiv các **sản xuất**
- Mỗi quy tắc là một dòng bao gồm 1 ký hiệu và 1 xâu được ngăn cách bởi dấu mũi tên
- Ký hiệu \equiv **biến** \equiv Các ký hiệu in hoa
- Ký hiệu kết thúc \equiv Các ký hiệu in thường, số hoặc ký tự đặc biệt
- **Biến ban đầu** thường xuất hiện bên trái của quy tắc trên cùng

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T \times F \mid F$$

$$F \rightarrow (E) \mid a$$

- Dẫn xuất

$$\begin{aligned} E &\Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow a + T \\ &\Rightarrow a + F \Rightarrow a + (E) \Rightarrow a + (T) \\ &\Rightarrow a + (T \times F) \Rightarrow a + (F \times F) \\ &\Rightarrow a + (a \times F) = a + (a \times a) \end{aligned}$$

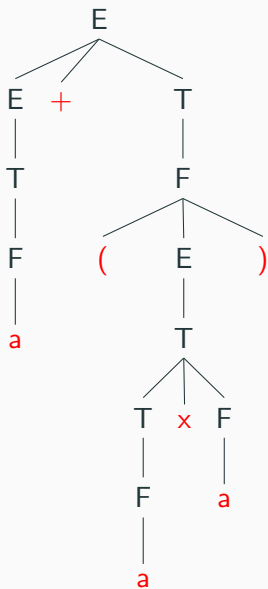
Cũng có thể viết:

$$E \xRightarrow{*} a + (a \times a)$$

$$E \xRightarrow{*} a + (E) \Rightarrow a + (T) \xRightarrow{*} a + (a \times a)$$

- Dẫn xuất trái nhất: Luôn lựa chọn dẫn xuất ở bên trái
 $\dots \Rightarrow F + T \Rightarrow a + T \Rightarrow \dots a + (a \times a)$
- Dẫn xuất phải nhất: Luôn lựa chọn dẫn xuất ở bên phải
 $\dots \Rightarrow F + T \Rightarrow F + F \Rightarrow \dots a + (a \times a)$

Cây dẫn xuất



Định nghĩa hình thức

$$\text{CFG: } \mathbf{G} = (\mathbf{V}, \Sigma, \mathbf{R}, \mathbf{S})$$

Trong đó:

- \mathbf{V} là tập hữu hạn gồm các **biến**
- Σ là tập hữu hạn các ký hiệu kết thúc $\Sigma \cap \mathbf{V}$
- \mathbf{R} tập các quy tắc
- \mathbf{S} biến bắt đầu