ĐẠI HỌC ĐÀ NẴNG

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN

CHƯƠNG TRÌNH DỊCH

ĐỀ TÀI :

Phân Tích Cú Pháp   
Bằng Phương Pháp Đệ Quy Không Quay Lui Cho Ngôn Ngữ C

GVHD : TS. NGUYỄN THỊ MINH HỶ

SINH VIÊN : NGUYỄN THỊ THƯ  
 NGUYỄN THỊ ANH THƯ

LỚP : 16TCLC3

**ĐÀ NẴNG, 2019**

# LỜI NÓI ĐẦU

Chương trình dịch là một môn học của ngành khoa học máy tính. Cùng với việc phát triển của các chuyên ngành: ngôn ngữ hình thức và otomat ,và các chuyên ngành khác, lý thuyết thiết kế biên dịch ngày càng một hoàn thiện hơn.

Ngày nay lý thuyết biên dịch đã được ứng dụng không những trong việc thiết kế trình biên dịch cho ngôn ngữ lập trình mà còn trong lĩnh vực trí tuệ nhân tạo-hiểu ngôn ngữ tự nhiên,dịch máy…

Đồ án môn học:Chương trình dịch là dịp tốt để sinh viên nắm vững lý thuyết và cách thiết kế trình biên dịch được tốt.

Chúng em đã thực hiện đồ án với đề tài 10:

**Phân tích cú pháp bằng phương pháp đệ qui không quay lui cho ngôn ngữ C**

Chúng em xin chân thành cảm ơn sự hướng dẫn tận tình của cô **Nguyễn Thị Minh Hỷ** và sự đóng góp ý kiến của các bạn,nhằm giúp cho đồ án được hoàn thành đúng yêu cầu. Mặc dù đã cố gắng tham khảo nhiều tài liệu nhưng trong quá trình thực hiện khó tránh khỏi sai sót và hạn chế. Rất mong nhận được ý kiến, đóng góp từ thầy cô và các bạn.

**Sinh viên thực hiện**

Nguyễn Thị Thư

Nguyễn Thị Anh Thư

**Lớp:16TCLC3**

MỤC LỤC

[LỜI NÓI ĐẦU 1](#_Toc10668885)

[**I.** **LÝ THUYẾT VỀ PHÂN TÍCH CÚ PHÁP BẰNG PHƯƠNG PHÁP ĐỆ QUY KHÔNG QUAY LUI** 3](#_Toc10668886)

[1. TÌM HIỂU CHUNG 3](#_Toc10668887)

[2. TRÌNH TỰ THỰC HIỆN 3](#_Toc10668888)

[a) Xây dựng văn phạm 3](#_Toc10668889)

[b) Xây dựng biểu đồ cú pháp 3](#_Toc10668890)

[c) Chương trình con biểu diễn cho biểu đồ cú pháp 3](#_Toc10668891)

[**II.** **ÁP DỤNG PHÂN TÍCH CÚ PHÁP CHO NGÔN NGỮ C** 7](#_Toc10668892)

[1) XÂY DỰNG VĂN PHẠM 7](#_Toc10668893)

[ Xây dựng văn phạm 7](#_Toc10668894)

[ Biến đổi văn phạm LL(1) 9](#_Toc10668895)

[2) XÂY DỰNG BIỂU ĐỒ CÚ PHÁP 9](#_Toc10668896)

[3) Xây dựng chương trình con phụ trợ 19](#_Toc10668897)

[**III.** **DEMO** 19](#_Toc10668898)

[TÀI LIỆU THAM KHẢO 24](#_Toc10668899)

1. **LÝ THUYẾT VỀ PHÂN TÍCH CÚ PHÁP BẰNG PHƯƠNG PHÁP ĐỆ QUY KHÔNG QUAY LUI**
2. TÌM HIỂU CHUNG

* Là phương pháp phân tích cú pháp từ trên xuống
* Về mặt nguyên lý giống phương pháp tiên đóan
* Khác về lập trình: không tra bảng tiên đoán M mà vẫn mô phỏng trực tiếp
* Thay stack bằng sự đệ quy trong chương trình
* Một kí hiệu chưa kết thúc :được biểu diễn bằng một biểu đồ cú pháp
* Một biểu đồ cú pháp: biểu diễn bằng 1 chương trình con

1. TRÌNH TỰ THỰC HIỆN
2. Xây dựng văn phạm

* Xây dựng văn phạm chính quy
* Biến đổi từ văn phạm chính quy sang văn phạm LL(1) theo 3 quy tắc

1. Xây dựng biểu đồ cú pháp

* Kí hiệu chưa kết thưc đặt :
* Kí hiệu kết thúc đặt :
* Ví dụ : E -> TE’

E’

T

1. Chương trình con biểu diễn cho biểu đồ cú pháp
2. Sự kết hợp liên tiếp của các nút : sự kế tiếp của các đoạn chương trình tương ứng

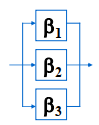
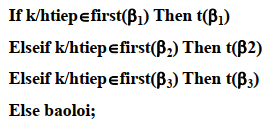
Ví dụ: E có đoạn chương trình t(E)

E’

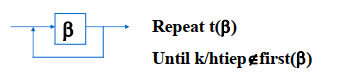
T

t(T); t(E’);

1. Sự rẽ nhánh tạo thành cấu trúc chọn

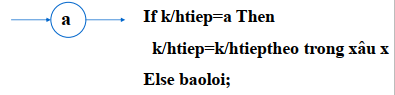
 

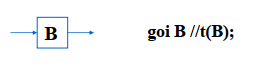
1. Lặp kiểm tra điều kiện sau



1. Lặp kiểm tra điều khiện trước



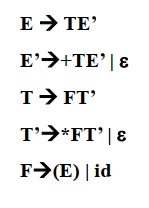




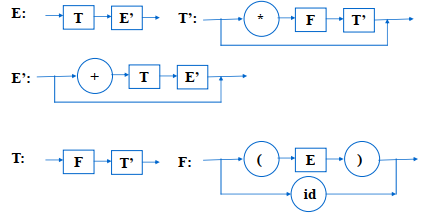
1. Thuật toán

* k/htiep: ký hiệu kết thúc
* function Dockh: trả về ký hiệu tiếp theo trong x
* Procedure Baoloi: cout lỗi, dừng
* Procedure B1: các chương trình con biểu diễn A∈Δ
* Procedure PTCP: đọc chuỗi x, nếu k/htiep=$ thì x đúng cú pháp else baoloi;

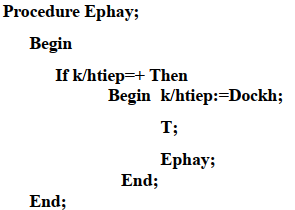
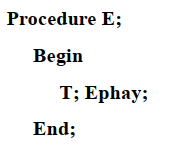
1. Ví dụ:

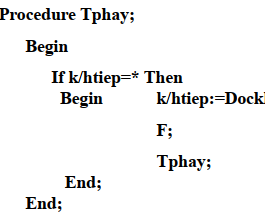
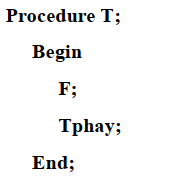


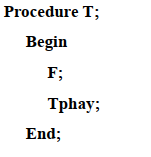
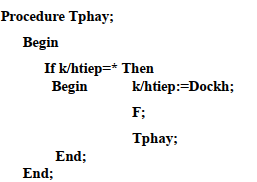
Biểu đồ cú pháp:

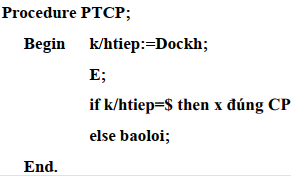
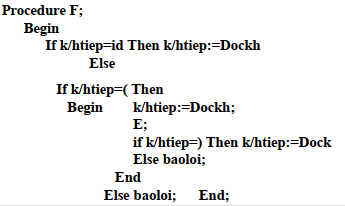


Chương trình con:







1. **ÁP DỤNG PHÂN TÍCH CÚ PHÁP CHO NGÔN NGỮ C**
2. XÂY DỰNG VĂN PHẠM

* Xây dựng văn phạm

G={ Σ, Δ,S,P } trong đó

Σ = token (từ phân tích từ vựng)

Δ={START,KB\_MAIN,KB’,BIEN\_HAM\_NMH,O,F,TV,KH,KH’,TENTV,TEN,KIEU,SO,STRING,STRING’,LK,B,MAIN,CT,X,HAM\_NMH,IN,OUT,IF,LOOP,STEP,OP,BT,BT’,P\_SO,SN,S,D,SO’,OP’,TTTH,TTTG,TTG,TTLG,A,K,P\_TOAN,P\_TG,P\_G,I,SW,C,D,BR,BR’,MANG,GAN,OP\_LOIGOIHAM, LOIGOIHAM, GT}

S={START}

P={

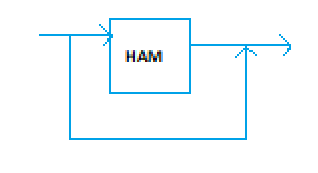
* START -> KB\_MAIN | KB\_MAIN HAM
* KB\_MAIN -> KB MAIN
* KB -> TV\_KH KB’
* KB’ -> KB | ε
* TV\_KH -> TV KH | TV
* TV -> #include<TENTV> |#define STRING SO| typedef STRING STRING;
* TENTV -> STRING
* STRING -> a STRING’|..|z|A|..|Z|0|..|9|%|/|? STRING’
* STRING’ -> STRING | ε
* SO -> 0 S’|..|9 S’
* S’ -> S | ε
* KH -> BIEN\_ HAM \_ NMH | HANG | KH’
* KH’ -> KH |ε
* BIEN\_HAM\_NHM -> KIEU TEN BHN
* BHN -> X ; | HAM\_NMH
* KIEU -> int|short|float|double|long|long long |unsigned int|unsigned char |bool|byte|char
* X -> MANG | LK | O
* LK -> LK’ B
* LK’ -> TEN
* TEN->a TEN’| b TEN’| c TEN’|… |\_ TEN’
* TEN’->a TEN’| b TEN’| c TEN’|… |\_ TEN’| 0 TEN’|…| 9 TEN’| ε
* B -> , LK | ε
* MANG ->[ SO ] GAN
* GAN -> ={PT}
* PT -> STRING PT\_PP
* PT\_PP -> , PT | ε
* O -> = STRING
* HAM\_NHM -> (STRING) F
* F -> NHM | HAM
* NHM -> ;
* HAM-> { CT}
* HANG -> const TEN SO
* MAIN -> main ( ){ CT}
* CT ->BIEN | IN |OUT| LOOP|STEP|IF |OP\_LOIGOIHAM|CT’
* CT’-> CT| ε
* IN -> scanf(“STRING”);
* OUT ->printf(“STRING”);
* LOOP -> while(BT){CT} | do {CT} while (BT); |for (BIEN; BT;OP) V
* V-> {CT} | CT
* BT -> STRING BT’ | (STRING BT’)
* BT’ -> P\_SO STRING | ε
* P\_SO -> < | > |=| != | >= | <= | ==
* STEP -> continue; | break;
* OP\_LOIGOIHAM-> TEN OP\_LOIGOIHAM’ ;
* OP\_LOIGOIHAM’-> OP’ | LOIGOIHAM
* LOIGOIHAM -> ( STRING)
* OP’ -> TTTH | TTTG | TTG| TTLG
* TTTH -> = A
* A -> GT P\_TOAN | (GT P\_TOAN ) P\_TOAN
* GT -> STRING
* P\_TOAN -> +A | -A| /A| % A| \*A| ε
* TTTG -> ++ | --
* TTG -> P\_G A
* P\_G -> += | -= | /= | \*= | %=
* TTLG -> BT K | (BT ) K
* K -> && TTLG | || TTLG | ! TTLG | ε
* IF -> I | SW
* I -> if (TTLG) I’
* I’ -> {CT} | CT I’’
* I’’ -> else I’ | ε
* SW -> switch (BT) { C D }
* C-> case (STRING) : CT BR
* BR -> BR’ C
* BR’-> break; | ε
* D -> default : CT |ε
* Biến đổi văn phạm LL(1)
* START -> KB\_MAIN | KB\_MAIN HAM
* START -> KB\_MAIN START’
* START’ -> HAM | ε
* TV\_KH -> TV KH | TV
* TV\_KH -> TV TV\_KH’
* TV\_KH’ -> KH | ε

1. XÂY DỰNG BIỂU ĐỒ CÚ PHÁP

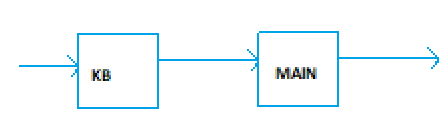
**START** :



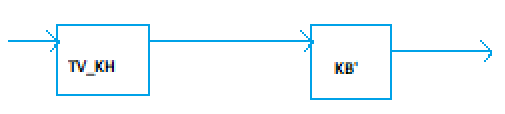
**START**’ :



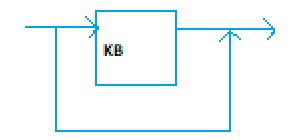
**KB\_MAIN:**



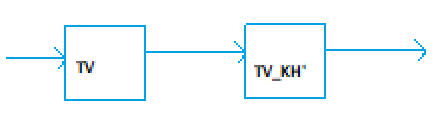
**KB:**



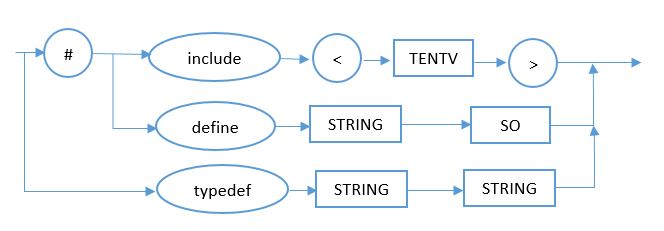
**KB’ :**



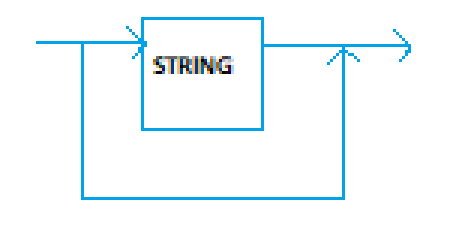
**TV\_KH :**



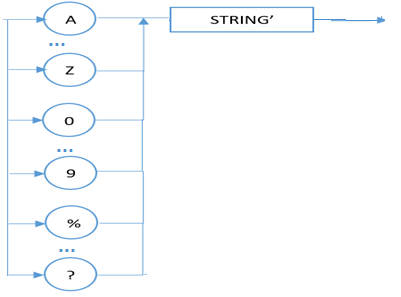
**TV:**



**TENTV:**



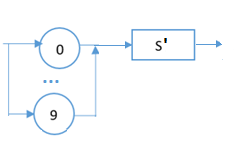
**STRING:**



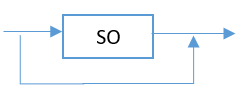
**STRING’:**



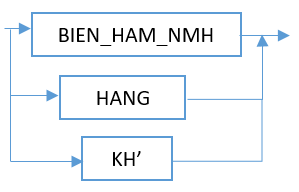
**SO:**



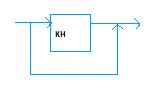
**S’:**



**KH:**



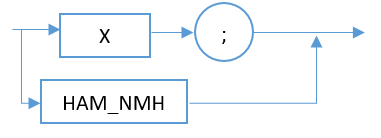
**KH’:**



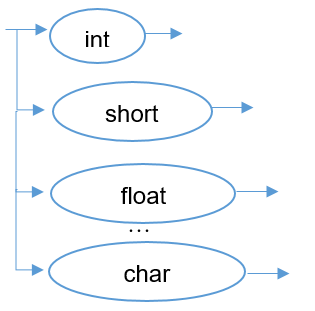
**BIEN\_HAM\_NMH:**



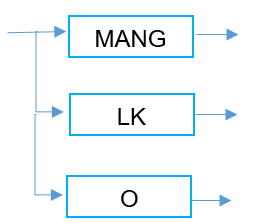
**BNH:**



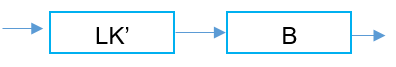
**KIEU:**

****

**X:**

****

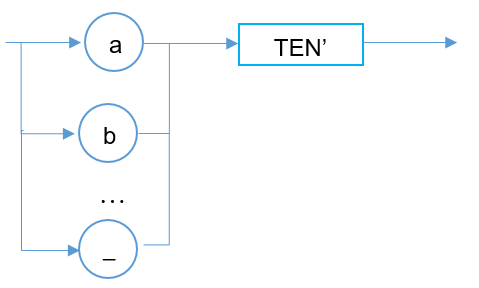
**LK:**

****

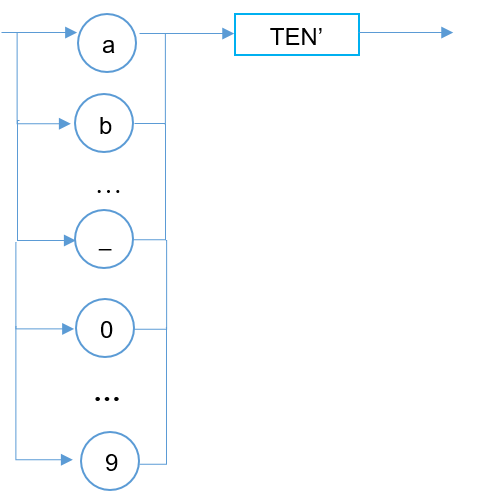
**LK’:**

****

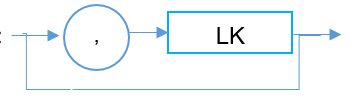
**TEN:**

****

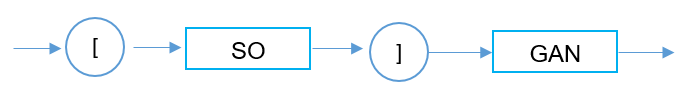
**TEN’:**

****

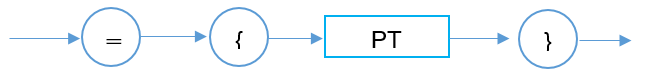
**B:**

****

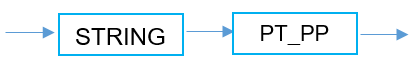
**MANG:**

****

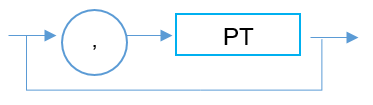
**GAN:**

****

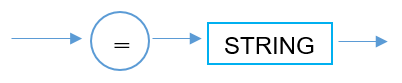
**PT:**

****

**PT\_PP:**

****

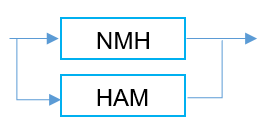
**O:**

****

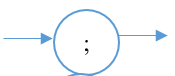
**HAM\_NMH:**



**F:**

****

**NMH:**

****

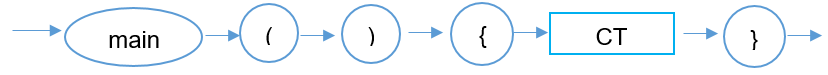
**HAM:**

****

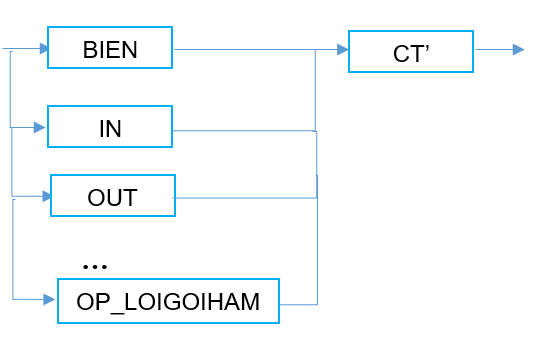
**HANG:**

****

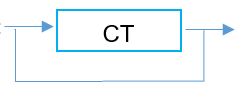
**MAIN:**

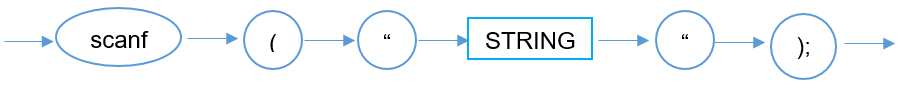
****

**CT:**

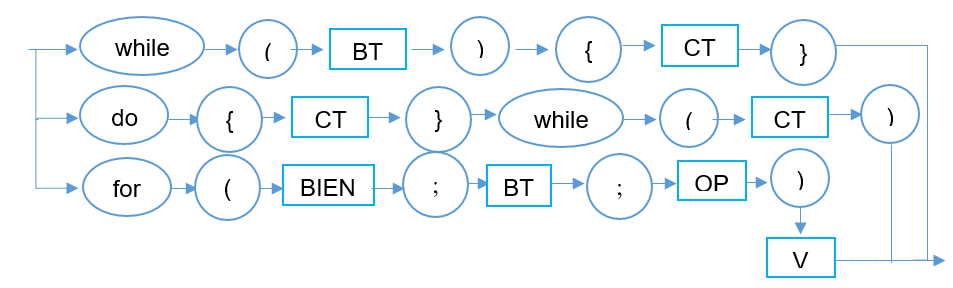
****

**CT’:**

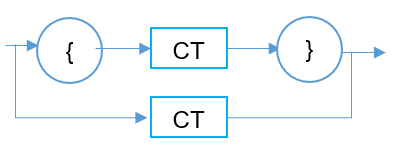
****

**IN: **

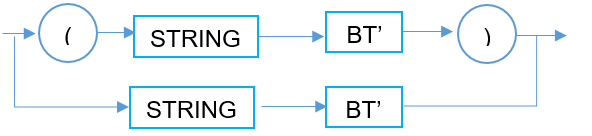
**OUT: **

**LOOP” **

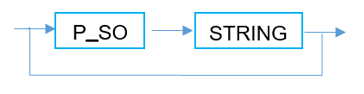
**V:**

****

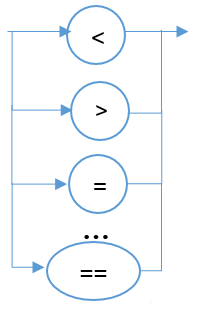
**BT:**

****

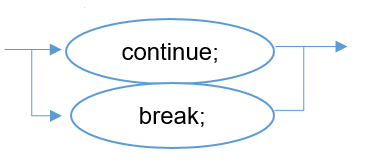
**BT’:**



**P\_SO:**

****

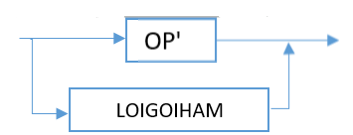
**STEP:**

****

**OP\_LOIGOIHAM:**



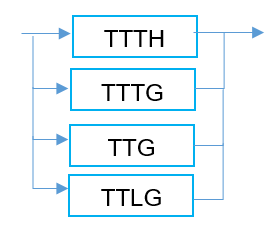
**OP\_LOIGOIHAM’:**



**LOIGOIHAM:**



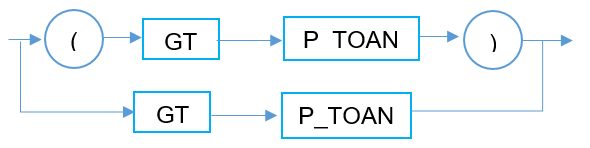
**OP’:**

****

**TTTH:**

****

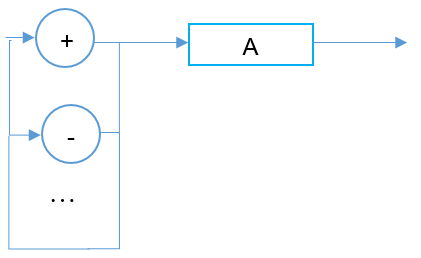
**A:**

****

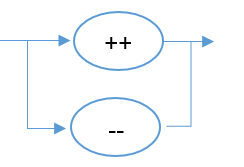
**GT:**



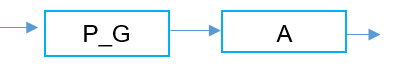
**P\_TOAN:**

****

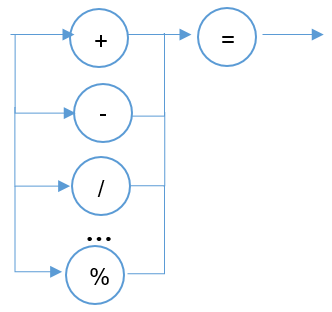
**TTTG:**

****

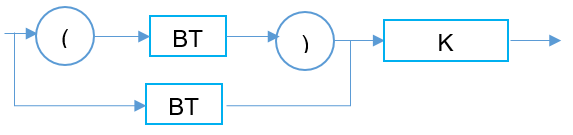
**TTG:**

****

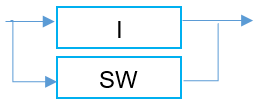
**P\_G:**

****

**V:**

****

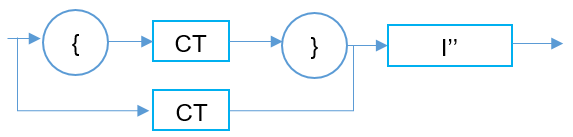
**IF:**

****

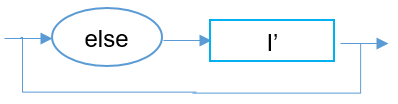
**I:**

****

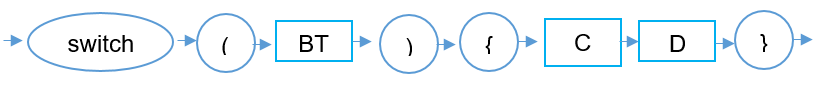
**I’:**

****

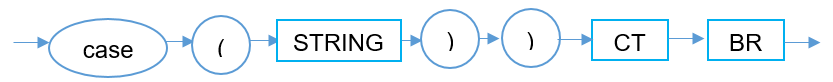
**I’’:**

****

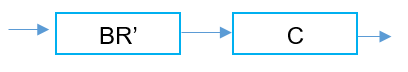
**SW:**

****

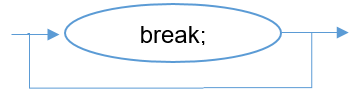
**C:**

****

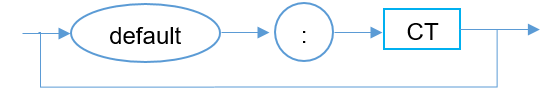
**BR:**

****

**BR’:**

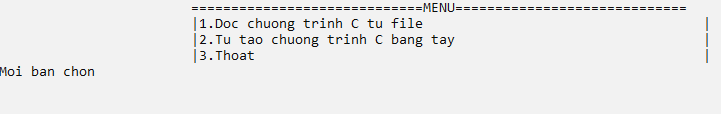
****

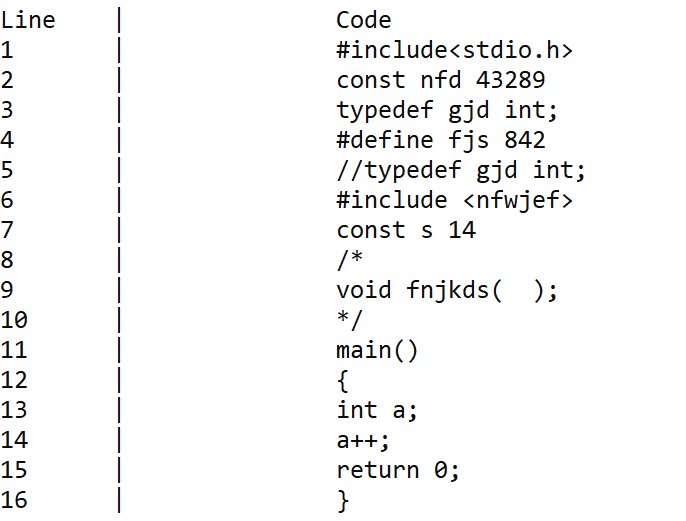
**D:**

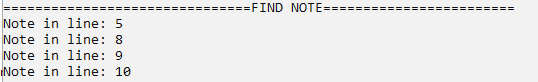


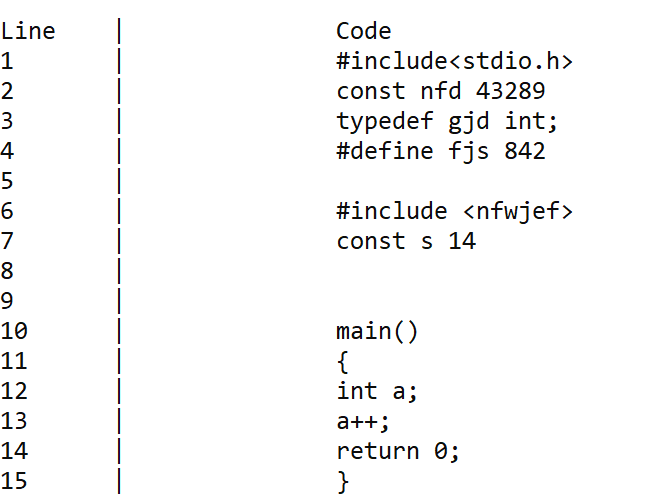
1. Xây dựng chương trình con phụ trợ
   * Đọc, ghi file, mở file mới để tạo chuỗi vào.
   * Xây dựng hàm
     1. Đọc kí tự kết thúc :char dockh();
     2. Báo lỗi thiếu kí tự, lỗi kí tự tại dòng nào đó: baoloi(), baoloithieu(kt)
     3. Kiểm tra chuỗi kí tự:void checkstring(string k);
     4. Đọc bỏ qua kí tự trắng(enter, tab..): checkc();
     5. Báo có chú thích trong chuỗi: Checkcmt();
     6. Lược bỏ chú thích trong chuỗi: XoaCMT();
     7. Các hàm con tương ứng với các sản xuất trên biểu đồ cú pháp
2. **DEMO**

**-**Giao diện bắt đầu:



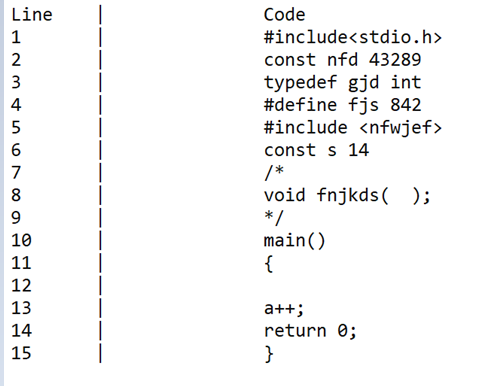
* Test case 1: có chú thích, không lỗi
  + Code
  + Phát hiện chú thích



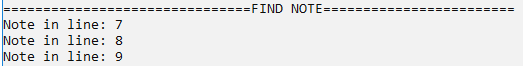
* + Sau khi lược bỏ chú thích
  + Kết quả



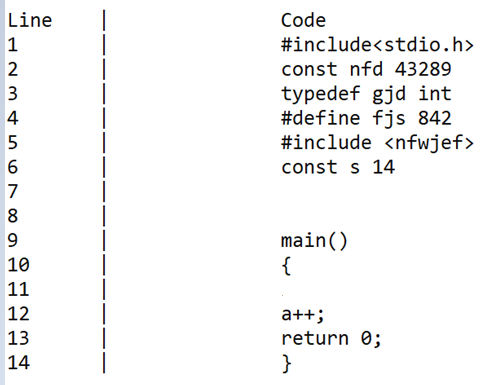
* Test case 2: có chú thích, có lỗi
  + Code



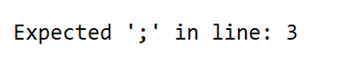
* + Phát hiện chú thích



* + Sau khi lược bỏ chú thích



* + Kết quả



TÀI LIỆU THAM KHẢO

[1]Nguyễn Thị Minh Hỷ, *Giáo trình Chương trình dịch,* Đại học Bách Khoa Đà Nẵng.