**KỸ THUẬT LẬP TRÌNH**

ĐỒ ÁN MÔN HỌC #1  
BÁO CÁO  
NGUỒN THAM KHẢO

Tên SV : Phạm Minh Chiến

MSSV : 1612052  
Môn : Kỹ thuật lập trình

Tên đồ án : **Đếm từ Tiếng việt**

C:\Users\tdqua_000\Dropbox\SS-Slides\DeCuong-CDIO\Template CDIO v4.2\Templates\Hinh anh\LogoTruong.png

Khoa Công nghệ Thông tin

Đại học Khoa học Tự nhiên TP HCM

Mục lục

[1. Tham khảo ý tưởng: 2](#_Toc484682711)

[ Cấu tạo từ: 2](#_Toc484682712)

[ Thuật toán tách từ: 2](#_Toc484682713)

[2. Xử lí Tiếng Việt trong C/C++: 2](#_Toc484682714)

[ Tham khảo wchar và cách encoding UTF-8, UTF-16: 2](#_Toc484682715)

[ Tham khảo về các lệnh wchar, wopenf: 2](#_Toc484682716)

[ Chuyển từ wchar\_ sang char: 2](#_Toc484682717)

[ Đoạn mã Unicode tiếng việt có dấu: 2](#_Toc484682718)

[3. Nguồn từ điển: 2](#_Toc484682719)

[ Phân mềm từ điển Bkav eDict: 2](#_Toc484682720)

# Tham khảo ý tưởng:

## Cấu tạo từ:

* [https://tiengvietmenyeu.wordpress.com/2011/12/30/cac-kieu-cau-tao-tu-tieng-viet](https://tiengvietmenyeu.wordpress.com/2011/12/30/cac-ki%E1%BB%83u-c%E1%BA%A5u-t%E1%BA%A1o-t%E1%BB%AB-ti%E1%BA%BFng-vi%E1%BB%87t/)
* <http://nguvan.hnue.edu.vn/Nghiencuu/Ngonngu/tabid/100/newstab/703/Default.aspx>
* <http://vietjack.com/soan-van-lop-6/tu-va-cau-tao-cua-tu-tieng-viet.jsp>
* <https://vi.wikipedia.org/wiki/T%E1%BB%AB#.C4.90.C6.A1n_v.E1.BB.8B_c.E1.BA.A5u_t.E1.BA.A1o>

## Thuật toán tách từ:

* <http://viet.jnlp.org/kien-thuc-co-ban-ve-xu-ly-ngon-ngu-tu-nhien/thuat-toan-tach-tu-tokenizer/thuat-toan-tach-tu>
* <http://www.ddth.com/archive/index.php/t-414400.html>
* <http://vietunicode.sourceforge.net/forum/viewtopic.php?t=241>
* <http://www.ngoquyen.edu.vn/thnguyentrai-ngoquyen/681/14679/37970/71506/suu-tam/kien-thuc-ve-tu-loai-tieng-viet-.aspx>

# Xử lí Tiếng Việt trong C/C++:

## Tham khảo wchar và cách encoding UTF-8, UTF-16:

* <https://daynhauhoc.com/t/hoi-ve-wchar-t-trong-c-go-tieng-viet-co-dau-trong-c/25100>
* [http://stackoverflow.com/questions/20195262](http://stackoverflow.com/questions/20195262/how-to-read-an-utf-8-encoded-file-containing-chinese-characters-and-output-them)

## Tham khảo về các lệnh wchar, wopenf:

* <https://msdn.microsoft.com/en-us/library/yeby3zcb.aspx>
* <http://www.cplusplus.com/reference/cwchar/>
* <http://www.cplusplus.com/reference/cwchar/wcscmp/>

## Chuyển từ wchar\_ sang char:

* <http://www.cplusplus.com/forum/beginner/28880/>
* <http://vncoding.net/2015/11/16/convert-wchar-sang-multibyte-va-nguoc-lai/>

## Đoạn mã Unicode tiếng việt có dấu:

* <http://vietunicode.sourceforge.net/charset/>
* <http://usefulwebtool.com/en/characters_vietnamese.php>

# Nguồn từ điển:

## Phân mềm từ điển Bkav eDict:

* Chỉnh sửa từ file từ điển: *ref/****vv30K.dict***
  + - * **Tác giả từ điển và bản quyền:**
* Cơ sở dữ liệu này được biên soạn bởi Hồ Ngọc Đức các thành viên của Dự án từ điển tiếng Việt miễn phí [(http://www.informatik.uni-leipzig.de/~duc/Dict/)](http://www.informatik.uni-leipzig.de/~duc/Dict/)
* Bản quyền (C) 1997-2003 Dự án từ điển tiếng Việt miễn phí
* Cho phép phối lại, sửa đổi theo các điều khoản của Giấy phép Công cộng GNU như được xuất bản bởi Tổ chức Phần mềm Tự do;