# **BẢNG PHÂN TÍCH GIẢI THUẬT CHUYỂN ĐỔI**

# **TỪ CHUYỂN TỰ LATIN SANG AKHAR THRAH**

# **(của phần mềm "Xalih Akhar Cham")**

1. **Bảng chữ cái chữ Chăm**

* Phụ âm:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chuyển tự EFEO | k | kh | g | gh | ng | nga |  |
| Chữ Chăm | k | A | g | G | z | Z |  |
| Chuyển tự EFEO |  |  |  |  |  |  |  |
| Chữ Chăm |  |  |  |  |  |  |  |

* Nguyên âm:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chuyển tự EFEO | a | i | u | e | ai | o |  |
| Chữ Chăm |  |  |  |  |  |  |  |

1. **Phần chương trình**
   1. *Qui trình chạy tổng quát của chương trình*

Chữ Chăm truyền thống

Mã ký tự chữ Chăm

Chuyển tự EFEO

Hướng phát triển của chương trình trong tương lại có thể hỗ trợ chuyển đổi nhiều kiểu chuyển tự Latin khác nhau sang nhiều font chữ Chăm khác nhau nên ta xây dựng thêm lớp trung gian là “Mã ký tự chữ Chăm”.

Một Mã ký tự chữ Chăm sẽ tương ứng với một ký tự của chữ Chăm truyền thống. Mã ký tự chữ Chăm có giá trị như bảng dưới đây.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chữ Chăm | k | A | g | G | z | Z |  |
| Mã ký tự chữ Chăm | 0 | 1 | 2 | 3 | 4 | 5 |  |

* 1. *Chuyển đổi từ chuyển tự EFEO sang Mã ký tự chữ Chăm:*

Một ký tự chuyển tự Latin có thể có nhiều ký tự chữ Chăm truyền thống. Ví dụ như ký tự chuyển tự Latin “i” có thể là nguyên âm Ik, vần I đi kèm với ký tự kết thúc câu, hay là vần I kết thúc câu, ngoài ra cũng có thể là takai Kiak. Như vậy 1 ký tự chuyển tự Latin “i” mà có thể là 1 trong 4 ký tự chữ Chăm trên.

Để giải quyết vấn đề trên, ta chia quá trình chuyển đổi từ chuyển tự Latin sang Mã ký tự chữ Chăm thành 2 bước:

+ Bước 1: chuyển đổi từ chuyển tự Latin sang Mã ký tự chữ Chăm đùng bảng tra.

+ Bước 2: Xử lý kết quả từ bước 1 cho các trường hợp: takai akhar, akhar mâtai, sap apaoh.

* + 1. *Chuyển đổi từ chuyển tự Latin sang Mã ký tự chữ Chăm dùng bảng tra*
* Phụ âm:
* Nguyên âm:
* Các trường hợp đặc biệt:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Chuyển tự Latin | ieng | ao | aom | aong | au | ém | iéng |
| Mã ký tự chữ Chăm | {TakaiKiak(69), TakaiThekPaohNgâk(65) } | { DarSa(57), TraohAw(74) } | { DarSa(57), TraohAwTutTakaiMâk(76) } | {} | {} | {} | {} |

* + 1. *Xử lý dữ liệu*
* Chuyển đổi dấu câu (takai akhar):

Dấu câu trong chữ Chăm lúc nào cũng đứng trước nguyên âm, vần. Dấu câu không bao giờ đứng trước chữ kết thúc câu (akhar mâtai). Do đó, nếu ký tự chuyển đổi là ký tự trong bảng sau và ký tự tiếp theo không phải là ký tự chữ kết thúc câu thì ta tiến hành chuyển đổi.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mã ký tự | Lak(29) | Rak(28) | Uk(37) | Ik(36) |
| Dấu câu (takai akhar) | Takai Klak(71) | Takai Krak(68) | Takai Kuak(70) | Takai Kiak(69) |

* Chuyển đổi dar sa, takai kuk:
* Xử lý ký tự kết thúc câu:

+ Chuyển đổi thành chữ kết thúc câu:

Nếu ký tự chuyển đổi là ký tự trong bảng sau và đứng ở cuối câu thì ta sẽ chuyển thành chữ kết thúc câu.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mã ký tự | Kak | Gak | .. | .. |
| Ký tự kết thúc câu (Akhar Mâtai) | Kak Mâtai () | Gak Mâtai | .. | .. |

+ Chuyển đổi vần kết thúc câu:

Nếu ký tự chuyển đổi là ký tự trong bảng sau và đứng ở cuối câu thì ta sẽ chuyển thành vần kết thúc câu.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mã ký tự | Ik() | É() | Takai Thek () | .. |
| Ký tự kết thúc câu (Akhar Mâtai) | {Takai kik tut takai mâk dalem} | {Dar sa, Balau tapong} | {Palau tapong ()} | .. |

Nếu

Kết quả trả về ở bước 1 chỉ là dữ liệu thô, đa phần chỉ đúng các ký tự đứng đầu câu. Do đó, ta phải chuyển đổi các ký tự kết thúc câu, lang li kuk, akhra mâtai, takai akhar,…

Ví dụ: từ “buk” khi qua bước 1 trả về array {22,37,0} tương ứng với {akhar Bak, akhar Uk, akhar kak}. Ta tiến hành chuyển đổi akhar Uk sang thành takai Uk, và akhar kak thành akhar Kak mâtai. Do logic xử lý ở bước này không rõ ràng, đa phần là chấp vá nên TNPTN không thể giải thích rõ hơn. Chi tiết giải thuật có thể tham khảo trong source code đính kèm. Như vậy kết quả cuối cùng sẽ là {22, 66, 41}.

* ***Chú ý***

Ngoài ra, một số từ có logic chuyển đổi nhập nhằng nên TNPTN đưa vào trường hợp đặc biệt. Việc chuyển đổi thực hiện trực tiếp, không thông qua 2 bước trên.

|  |  |
| --- | --- |
| **Chuỗi chuyển tự đặc biệt** | **Array of Key Code** |
| Xaai | {31,58,39} |
| Ai | … |
| Aia | … |
| Ia | … |
| Ppo | … |
| Liua | … |
| … | … |

Ví dụ: Chuổi ký tự “xaai” sẽ chuyển thành mảng của Key Code là {31,58,39}, trong đó 31 là akhar Xah, 58 là dar dua, 39 là akhar Ai.

## Key Code sang Font Cham

Ánh xạ từng ký tự key code trong array trả về từ mục 1.1 sang ký tự chữ Chăm của font tương ứng.

Ví dụ:

Chuyển đổi mã Key Code sang font Yapata. Chuổi chuyển tự là “buk”, sau khi qua mục 1.1 ta được array {22, 66, 41}. Sau khi ánh xạ ta được “b~K”. Khi hiện thị bằng font akhar Thrah 1 của Yapata sẽ là: b~K

|  |  |
| --- | --- |
| **Mã Key Code** | **Ký tự font Yapata** |
| 0 | k |
| 1 | A |
| 2 | g |
| 3 | G |
| 4 | z |
| 5 | Z |
| … | … |