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

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

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

* ***Giới thiệu chung***

Do yêu cầu của phần mềm phải có chức năng chuyển từ nhiều loại chuyển tự sang nhiều loại font Cham thông dụng hiện nay (multi to multi) nên nhóm Tanginpantangin (TPT) đã sử dụng một cấu trúc dữ liệu trung gian (Key Code).

Key Code

Font Yapata

Chuyển tự EFEO

Font BBSSCC

Chuyển tự Inrasara

Font GilaiPraung

Chuyển tự BBSSCC

Font EFEO

Font Cam Unicode

Như vậy đầu tiên ta sẽ chuyển chữ chuyển tự sang array của Key Code, sau đó sẽ chuyển array of Key Code sang các ký tự của font chữ Cham tương ứng.

Key Code đơn giản chỉ là enum.

|  |  |
| --- | --- |
| **Key Code [enum]** | |
| **Name** | **Value** |
| Kak | 0 |
| Khak | 1 |
| Gak | 2 |
| Ghak | 3 |
| Ngâk | 4 |
| Ngak | 5 |
|  | |
| Cak | 6 |
| Chak | 7 |
| Jak | 8 |
| Jhak | 9 |
| Nyâk | 10 |
| Nyak | 11 |
| Njak | 12 |
|  | |
| Tak | 13 |
| Thak | 14 |
| Dak | 15 |
| Dhak | 16 |
| Nâk | 17 |
| Nak | 18 |
| Ndak | 19 |
|  | |
| Pak | 20 |
| Phak | 21 |
| Bak | 22 |
| Bhak | 23 |
| Mâk | 24 |
| Mak | 25 |
| Mbak | 26 |
|  | |
| Yak | 27 |
| Rak | 28 |
| Lak | 29 |
| Wak | 30 |
| Xak | 31 |
| Hak | 32 |
|  | |
| PakPraong | 33 |
| SakPraong | 34 |
|  | |
| Ak | 35 |
| Ik | 36 |
| Uk | 37 |
| É | 38 |
| Ai | 39 |
| Ok | 40 |
|  | |
| KakMatai | 41 |
| GakMatai | 42 |
| NgâkMatai | 43 |
| CakMatai | 44 |
| TakMatai | 45 |
| NâkMatai | 46 |
| PakMatai | 47 |
| TutTakaiMâk | 48 |
| YakMatai | 49 |
| RakMatai | 50 |
| LakMatai | 51 |
| WakMatai | 52 |
| XakMatai | 53 |
| PaohDaNih | 54 |
|  | |
| Balau | 55 |
| BalauTapong | 56 |
| DarSa | 57 |
| DarDua | 58 |
| TakaiKik | 59 |
| TakaiKikTutTakaiMâkDalem | 60 |
| TakaiKikTutTakaiMâkLingiw | 61 |
| TakaiKikTutTakaiYak | 62 |
| TakaiThek | 63 |
| TakaiThekTutTakaiMâk | 64 |
| TakaiThekPaohNgâk | 65 |
| TakaiKuk | 66 |
| TakaiKâk | 67 |
| TakaiKrak | 68 |
| TakaiKiak | 69 |
| TakaiKuak | 70 |
| TakaiKlak | 71 |
| TakaiKlakTakaiKuak | 72 |
| TakaiKlakTakaiKuk | 73 |
| TraohAw | 74 |
| TraohAwPaohNgâk | 75 |
| TraohAwTutTakaiMâk | 76 |
| PaohNgâk | 77 |
|  | |
| Sa | 78 |
| Dua | 79 |
| Kluw | 80 |
| Ppak | 81 |
| Limâ | 82 |
| Nem | 83 |
| Tajuh | 84 |
| Dalipan | 85 |
| Salipan | 86 |
| Saoh | 87 |
|  | |
| KanaingSa | 88 |
| KanaingDua | 89 |
| KanaingKaokMâTi | 90 |
| Colon | 91 |
| ExclamationMark | 92 |
| QuestionMark | 93 |
| Minus | 94 |
| Patuk | 95 |
| Square | 96 |

## Chuyển tự sang Key Code.

Việc chuyển từ ký tự chuyển tự sang Key Code bao gồm 2 bước.

***Bước 1: Chuyển đổi lần 1 dựa vào bảng ánh xạ.***

|  |  |
| --- | --- |
| **Ký tự chuyển tự** | **Key code** |
| a | 35 |
| â | 67 |
| ai | 58 |
| ao | {57,74} |
| aom | {57,76} |
| aong | {57,75} |
| b | 22 |
| bh | 23 |
| c | 6 |
| ch | 7 |
| d | 15 |
| dh | 16 |
| e | 63 |
| é | 38 |

***Bước 2: Xử lý kết quả***

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 |
| … | … |