



Czech University
of Life Sciences Prague

English to Bangla Translator

Instructor's Name

Ing. Himesha Prabhakara Wijekoon

Course: Theoretical Informatics

Submitted By

Mia Sakil

Description of the Issue

The goal of this project is to create an English-to-Bangla character translator using Prolog. The project addresses the practical problem of transliterating English (phonetic) text into Bangla characters based on the Avro phonetic keyboard layout. This system is intended to help users write Bangla text using English letters as input, which is a common approach in digital communication in Bangladesh.

Bengali alphabet বাংলা বর্ণমালা বা লিপি					
অ	আ	ই	ঈ	উ	ঊ
ঋ	এ	ঐ	ও	ঔ	
ক	খ	গ	ঘ	ঙ	
চ	ছ	জ	ঝ	ঞ	
ট	ঠ	ড	ঢ	ণ	
ত	থ	দ	ধ	ন	
প	ফ	ব	ভ	ম	
য	র	ল			
শ	ষ	স	হ		
ড়	ঢ়	য়			
ৎ	ং	ঃ	্		

Brief Description of the Subject Theory

For solving this issue, the relevant theory involves computational linguistics and the study of formal grammars. The project could rely on:

- **Pattern Matching:** Using Prolog's facts and rules to map English characters to Bangla characters based on phonetic rules.
- **Recursive Programming:** Translating character-by-character using recursion to handle input strings of arbitrary lengths.
- **Backtracking:** Ensuring correct matches for input characters by iterating through facts and rules systematically.
- **Prolog Programming:** Prolog is well-suited for symbolic reasoning, making it ideal for language translation tasks where rules can be encoded for the mapping between source and target languages.

Task Analysis

Inputs and Outputs

- **Input:** The program accepts either a single English character or a list of English characters as input.
- **Output:** The program generates the corresponding Bangla character for a single input or a list of Bangla characters for a sequence of inputs, depending on the provided input type.

Facts for Translating English to Bangla

The program uses predefined facts for স্বরবর্ণ (Sworoborno), ব্যঞ্জনবর্ণ (Benjonborno), and the English alphabet to enable accurate character-based translation from English to Bangla.

- **স্বরবর্ণ (Sworoborno):** Facts are defined for Bangla vowels and their corresponding English representations. For example:
 - o => অ
 - a => আ
 - i => ই
- **ব্যঞ্জনবর্ণ (Benjonborno):** Facts are defined for Bangla consonants and their English equivalents. For example:
 - k => ক
 - kh => খ
 - g => গ

Rules for the Translating Program

- **Single Character Translation:**
This rule translates individual English characters into their corresponding Bangla characters using predefined mappings.
- **List of Characters Translation:**
This rule processes a list of English characters and translates each character into Bangla, ensuring compound characters and sequences are handled appropriately.
- **Error Handling:**
This rule detects and handles errors, such as unsupported characters or invalid inputs, by providing meaningful messages or default behaviors.

Implementation Solution Description

Approaches Considered

1. Pure Fact-Based Translation:

- Define individual `char_translation` facts for all English-to-Bangla mappings.
- Use recursive rules to process character lists.
- Pros: Straightforward and easy to implement.
- Cons: Error handling and complex mappings (like diacritics) require additional effort.

2. Automata-Based Approach:

- Define a finite state automaton (FSA) with states representing transitions between vowels, consonants, and errors.
- Use transition rules for translating the input.
- Pros: Scalable and structured for complex rules.
- Cons: Overkill for simple transliteration tasks.

Selected Approach: The **fact-based translation** approach was chosen due to its simplicity and suitability for small-scale transliteration tasks.

Implementation: A Commented Programme

```
% স্বরবর্ণ sworoborno (vowels)
char_translation('o', 'অ').
char_translation('a', 'আ').
char_translation('i', 'ই').
char_translation('I', 'ঈ').
char_translation('u', 'উ').
char_translation('U', 'ঊ').
char_translation('rri', 'ঋ').
char_translation('e', 'এ').
char_translation('OI', 'ঐ').
char_translation('O', 'ও').
char_translation('OU', 'ঔ').

% ব্যঞ্জনবর্ণ benjonborno (Consonants)
char_translation('k', 'ক').
char_translation('kh', 'খ').
char_translation('g', 'গ').
char_translation('gh', 'ঘ').
char_translation('Ng', 'ঙ').
char_translation('c', 'চ').
char_translation('ch', 'ছ').
char_translation('j', 'জ').
char_translation('jh', 'ঝ').
char_translation('NG', 'ঞ').
char_translation('T', 'ট').
char_translation('Th', 'ঠ').
char_translation('D', 'ড').
char_translation('Dh', 'ঢ').
char_translation('N', 'ণ').
char_translation('t', 'ত').
char_translation('th', 'থ').
char_translation('d', 'দ').
char_translation('dh', 'ধ').
char_translation('n', 'ন').
char_translation('p', 'প').
char_translation('f', 'ফ').
char_translation('b', 'ব').
char_translation('v', 'ভ').
char_translation('m', 'ম').
char_translation('z', 'য').
char_translation('r', 'র').
```

```

char_translation('l', 'ল').
char_translation('sh', 'শ').
char_translation('Sh', 'ষ').
char_translation('s', 'স').
char_translation('h', 'হ').
char_translation('R', 'ড়').
char_translation('Rh', 'ঢ়').
char_translation('y', 'য়').
char_translation('t`', 'ত').
char_translation('ng', 'ং').
char_translation(':', 'ঃ').
char_translation('^', 'ূ').

% Rule to translate a single character
translate_char(EnglishChar, BanglaChar) :-
    char_translation(EnglishChar, BanglaChar).

% Error handling for unrecognized facts
translate_char(_, 'Not in facts').

% Rule to Translate a list of characters
translate_chars([], []).
translate_chars([E|Es], [B|Bs]) :-
    translate_char(E, B),
    translate_chars(Es, Bs).

```

Programme Description

Syntax and Semantics of Programme Inputs:

- **Input:** The program takes a English character (tokens) as input, such as “a”.
- **Output:** The program outputs the corresponding Bengali character, such as “আ”.

The input to the program could be a list of character, and the output would be the list of translated character.

Description of Programme Outputs with Examples:

- **Example 1:**
 - Input: ['o' , 'a' , 'i']
 - Output: ['অ' , 'আ' , 'ই']
- **Example 2:**
 - Input: ['k' , 'kh' , 'g']
 - Output: ['ক' , 'খ' , 'গ']