

### Online found effective cyclic Hash function:

FUNCTION hashCode(word):

  // Constants

  GOLDEN ← 0x9e3779b97f4a7c15     // 64-bit golden ratio

  ROT ← 7                           // bits to rotate each step

  BITS ← size\_in\_bits(unsigned long)

  seed ← 0

  FOR each character c IN word DO

    // 1) cyclic-shift left by ROT bits

    seed ← (seed << ROT) OR (seed >> (BITS – ROT))

    // 2) mix in the byte plus “golden” and shifts

    mixture ← toUnsignedLong(c)

      + GOLDEN

      + (seed << 6)

      + (seed >> 2)

    seed ← seed XOR mixture

  END FOR

  RETURN seed

```
=====
Size of HashTable           = 812739
Total Number of Collisions  = 1485947
Avg. Number of Collisions/Entry = 1.8
=====
find <word>                  : Search a word and its meanings in the dictionary.
import <path>                : Import a dictionary file.
add <word:meaning(s):language> : Add a word and/or its meanings (separated by ;) to the dictionary.
delTranslation <word:language> : Delete a specific translation of a word from the dictionary.
delMeaning <word:meaning:language> : Delete only a specific meaning of a word from the dictionary.
delWord <word>               : Delete a word and its all translations from the dictionary.
export <language:filename>   : Export a a given language dictionary to a file.
exit                        : Exit the program
```

```

>add UAE:Vereinigte Arabische Emirate:German
>add UAE:Emirats arabes unis:French
uae found in the Dictionary after 3 comparisons.
german : Vereinigte Arabische Emirate
french : Emirats arabes unis
>find success
success found in the Dictionary after 4 comparisons.
german : Erfolg; Gelingen; Sukzess; Sukzeß; Triumph; Ausgang; Folge; Gedeihen; guter Ausgang; Erfolge
spanish : éxito
french : réussite; succès
>find hello
hello found in the Dictionary after 0 comparisons.
german : Hallo
spanish : Cómo es'; Hola
french : Salut; Bonjour
>find dictionary
dictionary found in the Dictionary after 0 comparisons.
german : Dictionnaire; Diktionär; Lexikon; Wörterbuch
spanish : diccionario
french : dictionnaire
>find self
self found in the Dictionary after 1 comparisons.
german : eigene Person; Ich; selber; selbst
spanish : mismo
french : même

```

Compared to the normal polynomial hash function and the normal cyclic hash function we can see that this hash function is best out of all. This method is a bit more computationally intensive because it involves multiple shifts, a rotation, and an XOR operation per character. However, all that extra work pays off in terms of collision reduction. In our tests, this function performed the best, reducing collisions by up to two-thirds compared to the basic cyclic hash. The added complexity ensures that even minor changes in the input result in significantly different hash values, making it the most effective of the three functions we tested.

### **Polynomial Hash function:**

FUNCTION hashCode(key):

// A small prime base for the polynomial

$p \leftarrow 31$

// Initialize the rolling hash

hash  $\leftarrow 0$

// For each character in the input string

FOR each character c IN key DO

// Multiply the current hash by p, then add the byte value of c

hash  $\leftarrow$  hash \* p + toUnsignedInt(c)

END FOR

RETURN hash

END FUNCTION

```

812739 German words imported successfully.
=====
Size of HashTable          = 812739
Total Number of Collisions = 1481060
Avg. Number of Collisions/Entry = 1.8
=====
find <word>                : Search a word and its meanings in the dictionary.
import <path>              : Import a dictionary file.
add <word:meaning(s):language> : Add a word and/or its meanings (separated by ;) to the dictionary.
delTranslation <word:language> : Delete a specific translation of a word from the dictionary.
delMeaning <word:meaning:language> : Delete only a specific meaning of a word from the dictionary.
delWord <word>             : Delete a word and its all translations from the dictionary.
export <language:filename>  : Export a a given language dictionary to a file.
exit                      : Exit the program

>add UAE:Vereinigte Arabische Emirate:German
>add UAE:Emirats arabes unis:French
>find UAE
uae found in the Dictionary after 11 comparisons.
german : Vereinigte Arabische Emirate
french : Emirats arabes unis
>find success
success found in the Dictionary after 1 comparisons.
german : Erfolg; Gelingen; Sukzess; Sukzeß; Triumph; Ausgang; Folge; Gedeihen; guter Ausgang; Erfolge
spanish : éxito
french : réussite; succès
>find hello
hello found in the Dictionary after 0 comparisons.
german : Hallo
spanish : Cómo es'; Hola
french : Salut; Bonjour
>find dictionary
dictionary found in the Dictionary after 1 comparisons.
german : Dictionnaire; Diktionär; Lexikon; Wörterbuch
spanish : diccionario
french : dictionnaire
>find self
self found in the Dictionary after 0 comparisons.
german : eigene Person; Ich; selber; selbst
spanish : mismo
french : même

```

The constant  $p$  is usually a small prime number, like 31. This method is fast because it only requires a basic multiplication and addition for each character. However, the downside is that similar keys (like words with slight variations) tend to produce similar hash values, leading to more collisions. In our experiments, the polynomial hash function had the highest collision rate. Its simplicity comes at the cost of less effective data distribution, making it less ideal for datasets with closely related strings.

### Cyclic Hash Function:

FUNCTION hashCode(word):

// Initialize state

hash  $\leftarrow$  0

// Constants

$R \leftarrow 7$  // rotation amount

$\text{BITS} \leftarrow \text{size\_in\_bits}(\text{hash})$  // e.g. 64 on a 64-bit machine

```

// Process each byte of the input string
FOR each character c IN word DO
    // 1) Rotate the current hash value left by R bits
    // (bits that overflow on the left wrap around to the right)
    hash ← (hash << R) OR (hash >> (BITS – R))

    // 2) Mix in the next byte via XOR
    hash ← hash XOR toUnsignedInt(c)
END FOR

RETURN hash
END FUNCTION

```

```

812739 German words imported successfully.
=====
Size of HashTable           = 812739
Total Number of Collisions  = 1490089
Avg. Number of Collisions/Entry = 1.8
=====
find <word>                  : Search a word and its meanings in the dictionary.
import <path>                : Import a dictionary file.
add <word:meaning(s):language> : Add a word and/or its meanings (separated by ;) to the dictionary.
delTranslation <word:language> : Delete a specific translation of a word from the dictionary.
delMeaning <word:meaning:language> : Delete only a specific meaning of a word from the dictionary.
delWord <word>               : Delete a word and its all translations from the dictionary.
export <language:filename>    : Export a a given language dictionary to a file.
exit                         : Exit the program

```

```

>add UAE:Vereinigte Arabische Emirate:German
>add UAE:Emirats arabes unis:French
>find UAE
uae found in the Dictionary after 6 comparisons.
german : Vereinigte Arabische Emirate
french : Emirats arabes unis
>find success
success found in the Dictionary after 0 comparisons.
german : Erfolg; Gelingen; Sukzess; Sukzeß; Triumph; Ausgang; Folge; Gedeihen; guter Ausgang; Erfolge
spanish : éxito
french : réussite; succès
>find hello
hello found in the Dictionary after 1 comparisons.
german : Hallo
spanish : Cómo es'; Hola
french : Salut; Bonjour
>find dictionary
dictionary found in the Dictionary after 0 comparisons.
german : Dictionnaire; Diktionär; Lexikon; Wörterbuch
spanish : diccionario
french : dictionnaire
>find self
self found in the Dictionary after 1 comparisons.
german : eigene Person; Ich; selber; selbst
spanish : mismo
french : même

```

Here, R is a fixed rotation value, often 7. The bit rotation helps mix up the bits more thoroughly, making it harder for similar keys to produce similar hashes. This function requires a bit more computational effort than the polynomial hash, but the payoff is worth it: in our testing, the cyclic hash had fewer collisions than the polynomial hash. It's a good middle ground—still relatively simple but with noticeably better data distribution.