

Improving the Like Operator in VB.NET: LikeOperatorB

Introduction

The Like operator in VB.NET is widely used for pattern matching in strings. However, it has some inconsistencies, especially when handling diacritics and string expansion.

This article presents LikeOperatorB, a revised version of the LikeString operator, optimized for better performance and improved string comparison accuracy.

Key Improvements in LikeOperatorB

Several important modifications have been made to optimize the Like operator by removing unnecessary dependencies and enhancing Unicode support:

1 Removal of Ligatures and String Expansion Code

All code related to ligatures and string expansion has been removed, as it introduced unnecessary complexity.

This improves efficiency by avoiding additional comparisons that could affect performance.

2 Elimination of Char() Functions

The original LikeString used Char() to manipulate characters, introducing overhead and reliance on CompareInfo.

These functions have been removed to allow for more efficient handling.

3 Use of Text.Rune for String Extraction and Comparison

Instead of Char(), LikeOperatorB utilizes Text.Rune, which extracts and compares Unicode characters more effectively.

This resolves issues with variable-width characters and improves accuracy for different languages.

4 Added Documentation and Comments in the Code

Notes have been added to the source code to indicate which parts were modified.

5 New Option to Normalize Diacritics (Optional Feature)

A new option allows diacritics to be ignored in comparisons.

If enabled, characters like "ü" and "u" will be considered equivalent, making text matching more flexible.

If disabled, "Müller" and "Muller" will still be treated as different.

Usage Example

To use LikeOperatorB, simply replace Like with LikeOperatorB.LikeString:

Dim result As Boolean

result = LikeOperatorB.LikeString("éclair", "é*", CompareMethod.Text) ' Returns True

To ignore diacritics in comparisons:

Dim result As Boolean

**result = LikeOperatorB.LikeString("pingüino", "pinguino", CompareMethod.Text,
ignoreDiacritics:=True) ' Returns True**

Test Results

Extensive testing has been conducted on LikeOperatorB to evaluate its performance compared to the original Like operator.

The tests include:

- ✓ Binary comparisons (CompareMethod.Binary)
- ✓ Textual comparisons (CompareMethod.Text)
- ✓ Tests with international characters (ß, Œ, æ, éclair, 日本, etc.)
- ✓ Pattern matching with wildcards (*, ?, [])
- ✓ Comparisons with and without diacritics