**Регулярні вирази**

*System.Text.RegularExpressions.*

Класи простору імен System.Text.RegularExpressions

|  |  |
| --- | --- |
| Capture | Надає результати з окремого запису частини виразу. Capture представляє один підрядок для окремого успішного запису. |
| CaptureCollection | Представляє послідовність підряджання запису. CaptureCollection повертає набір записів, зроблених однією групою запису. |
| Group | Group представляє результати окремої групи запису. Група запису може записати нуль, один або більше рядків в один збіг через квантори, таким чином Group надає колекцію об'єктів Capture. |
| GroupCollection | Представляє колекцію груп запису. GroupCollection повертає набір груп запису в одному збігу. |
| Match | Представляє результати з окремого збігу регулярного виразу. |
| MatchCollection | Набір успішних збігів, виявлених шляхом ітеративного застосування шаблону регулярного виразу до вхідного рядка. |
| Regex | Постійне регулярне вираження. |
| RegexCompilationInfo | Надає відомості про регулярне вираження, що використовуються для компіляції регулярного виразу в окреме складання. |
| RegexRunner | Клас RegexRunner є базовим класом для скомпільованих регулярних виразів. |
| RegexRunnerFactory | Створює клас RegexRunner |

Ключовий клас, який дозволяє створювати регулярні вирази Regex – має кілька конструкторів.

1) новий Regex(string pattern)

*2) новий Regex(string pattern, RegexOptions options)*

де: pattern - це маска регулярного вираження;

*options*- Додаткові параметри, перерахування типу RegexOptions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Метасимволи, які використовуються в регулярних виразах C#* | | | | |
| **Символ** | **Значення** | | **приклад** | **Відповідає** |
| Класи символів | | | | |
| [...] | Будь-який із символів, вказаних у дужках | | [az] | У вихідному рядку може бути будь-який символ англійського алфавіту в нижньому регістрі |
| [^...] | Будь-який із символів, не вказаних у дужках | | [^0-9] | У вихідному рядку може бути будь-який символ, окрім цифр |
| . | Будь-який символ, окрім перекладу рядка або іншого роздільника Unicode-рядка | |  |  |
| \w | Будь-який текстовий символ, який не є пробілом, символом табуляції тощо. | |  |  |
| \W | Будь-який символ, який не є текстовим символом | |  |  |
| \s | Будь-який символ пробілу з набору Unicode | |  |  |
| \S | Будь-який непробілний символ із набору Unicode. Зверніть увагу, що символи \w і \S - це не те саме | |  |  |
| \d | Будь-які ASCII-цифри. Еквівалентно [0-9] | |  |  |
| \D | Будь-який символ відмінний від ASCII-цифр. Еквівалентно [^0-9] | |  |  |
| **Символи повторення** | | | | |
| {n, m} | Відповідає попередньому шаблону, повтореному не менше n і не більше m разів | | s{2,4} | "Press", "ssl", "progressss" |
| {n,} | Відповідає попередньому шаблону, повтореному n або більше разів | | s{1,} | "ssl" |
| {n} | Відповідає в точності n екземплярам попереднього шаблону | | s{2} | "Press", "ssl", але не "progressss" |
| ? | Відповідає нулю або одному примірнику попереднього шаблону; попередній шаблон є необов'язковим | | Еквівалентно {0,1} |  |
| + | Відповідає одному або більше екземплярам попереднього шаблону | | Еквівалентно {1,} |  |
| \* | Відповідає нулю або більше екземплярам попереднього шаблону | | Еквівалентно {0,} |  |
| **Символи регулярних виразів вибору** | | | | |
| | | Відповідає або подвыражению зліва, або подвыражению справа (аналог логічної операції АБО). | |  |  |
| (...) | Угруповання. Групує елементи в єдине ціле, яке можна використовувати з символами \*, +, ?, | і т.п. Також запам'ятовує символи, які відповідають цій групі для використання у наступних посиланнях. | |  |  |
| (?:...) | Лише угруповання. Групує елементи у єдине ціле, але не запам'ятовує символи, що відповідають цій групі. | |  |  |
| **Якірні символи регулярних виразів** | | | | |
| ^ | Відповідає початку рядкового виразу або початку рядка при багатостроковому пошуку. | | ^Hello | " Hello, world " , але з " Ok, Hello world " т.к. у цьому рядку слово "Hello" знаходиться не на початку |
| $ | Відповідає кінцю рядкового виразу або кінцю рядка при багатостроковому пошуку. | | Hello$ | "World, Hello" |
| \b | відповідає межі слова, тобто. відповідає позиції між символом \w та символом \W або між символом \w та початком або кінцем рядка. | | \b(my)\b | У рядку "Hello my world" вибере слово "my" |
| \B | Відповідає позиції, яка не є межею слів. | | \B(ld)\b | Відповідність знайдеться у слові "World", але не в слові "ld" |
| *Структура перерахування RegexOptions* | | | | |
| **Член** | | **Опис** | | |
| CultureInvariant | | Наказує ігнорувати національні установки рядка | | |
| ExplicitCapture | | Модифікує спосіб пошуку відповідності, забезпечуючи лише буквальну відповідність | | |
| IgnoreCase | | Ігнорує регістр символів у вхідному рядку | | |
| IgnorePatternWhitespace | | Видаляє з рядка пробіли, що не захищені керуючими символами, і дозволяє коментарі, що починаються зі знака фунта або хеша. | | |
| Multiline | | Змінює значення символів ^ і $ так, що вони застосовуються до початку та кінця кожного рядка, а не тільки до початку та кінця всього вхідного тексту | | |
| RightToLeft | | Наказує читати вхідний рядок праворуч наліво замість напрямку за замовчуванням — зліва направо (що зручно для деяких азіатських та інших мов, які читаються у такому напрямку) | | |
| Singleline | | Специфікує однорядковий режим, у якому точка (.) символізує відповідність будь-якому символу | | |