1. **RegularExpression**

### RegularExpression (Biểu thức chính qui)

Dùng để so khớp một chuỗi với một mẫu xem chuỗi có thỏa theo chuẩn của mẫu hay không.

Vd: so 1 chuỗi có phải là email hay không, so 1 dãy xem có phải là số xe TPHCM không.

Muốn xây dựng mẫu phải dựa vào các kí tự, mỗi kí tự có ý nghĩa riêng

Một số kí tự thường dùng

|  |  |
| --- | --- |
| **Kí tự** | **Mô tả** |
| . | Đại diện cho 1 kí tự bất kì trừ kí tự xuống dòng \n |
| \d | Kí tự chữ số tương đương [0-9] (một kí số) |
| \D | Kí tự không phải một kí số |
| \s | Đại diện cho kí tự trắng gồm cả tab, xuống dòng, về đầu dòng .Tương đương [\f\n\r\t\v] |
| \S | Kí tự không phải khoảng trắng [^\f\n\r\t\v] |
| \w | Kí tự chữ cái, chữ số, dấu \_ . Tương đượng [a-zA-Z0-9\_] |
| \W | Tương đương [^a-zA-Z0-9\_] |
| ^ | Bắt đầu 1 chuỗi hay 1 dòng |
| $ | Kết thúc 1 chuỗi hay 1 dòng |
| | | Cái này hoặc cái kia . Tương đương với or |
| [] | Đại diện 1 kí tự được liệt kê trong [] |
| [^] | Đại diện 1 kí tự không được liệt kê trong [] |
| () | Phân nhóm vd: ((a(b))c) sẽ khớp với b, ab, abc |
| ? | Tương đương {0,1}. Đứng trước từ 0 hay 1 lần. vd: A?B khớp với B hay AB |
| \* | Tương đương {0,} . Đứng trước từ 0 hoặc nhiều lần. A\*B khớp với B, AB, AAB, AAAB…. |
| + | Tương đương {1,}. Đứng trước từ 1 lần trở lên. A+B khớp với AB, AAB, AAAB…… |
| {n} | Kí tự trước nó xuất hiện đúng n lần |
| {min,max} | Kí tự trước nó xuất hiện ít nhấn min lần nhiều nhất max lần |

Ví dụ:

* Email đơn giản:

String regex=”\\w+@\\w+\\.\\w+”;

* Số CMND:

String regex=”\\d{10}”;

* Số điện thoại di động:

String regex=”0\\d{8,9}”;

* Số xe máy TPHCM

String regex =”5\\d-[A-Z]\\d-\\d{3}.{2}”;

* Địa chỉ web:

String regex=”<http://www.(\\w+\\.)\\.+\\w+>”;

Khi sử dụng ta phải dùng với class Pattern và class Matcher như ví dụ bên dưới

