

Appendix 4 – Some useful wildcard expressions

First let me try to explain my terminology:

1) Here is an example of what I call an ‘**expression**’:

`[0-9]{4}`

If you type this into the Find box and click ‘Use wildcard’ and then ‘Find Next’, it will jump to the next year, e.g. ‘2009’, well, to the next four-digit number.

2) Ordinary finds and expressions can be **grouped**, by using (), then those **groups** can be referred to by using `\1` for the first group, `\2` for the second, etc.

3) Using [] specifies a **range**, i.e. “Find **one** of these characters.”. For example...

`[02468]` finds an even number

`[0-9]` finds one digit

`[a-z]` finds a single lowercase character

`[A-Z]` finds a single uppercase character

Using `@` after a character or a range means “Find *one or more* of these – but find as *few* as possible.”

Using, say, `{2}` after a character or a range means “Find *two* of these.”

Using, say, `{2,4}` after a character or a range means “Find *two, three or four* of these.”

Using, say, `{2,}` after a character or a range means “Find *two, or more* of these – but find as *MANY* as possible.”

Examples:

`o{1,}` Finds: **profit**, **proof**reader, **spooof!**

`o@` Finds: **profit**, **proof**reader, **spooof!**

`o@f` Finds: **profit**, **proof**reader, **spooof!**

Useful example expressions:

`[a-z]{1,}` lowercase word

`[A-Z]{1,}` uppercase word

`[a-zA-Z]{1,}` mixed-case word, e.g. Beverley, McTaggart, BBC, hello

N.B. Some websites and books advocate use of `[A-z]`. This is **wrong**, (a) it does not find accented characters, such as in ‘façade’ or ‘déjà vu’, and (b) it **includes** the characters with ASCII codes between ‘Z’ and ‘a’, that is: `[, \,], ^, _`, and ``` (back tick).

`[0-9]{3,}` multi-digit number, for 100,000, it would first find just 100, then 000

`[,0-9]{5,}` multi-digit number with commas, e.g. 100,000 or 4567 or 1234,6785,000

`[hw][tw][tw][a-zA-Z0-9:/\._=?&]{6,}` URL (if you think of more characters needed, please tell me!)

`[A-Ω]` uppercase Greek character

- [α-ω] lowercase Greek character
- [А-Я] uppercase Cyrillic character (may be wrong – I don't know Cyrillic!)
- [а-я] lowercase Cyrillic character (may be wrong – I don't know Cyrillic!)
- [a-z]@ing present participles
- [a-z]@ed> past participles

N.B. `[a-z]{1,}ing` wouldn't work because for, say, 'running', the `[a-z]{1,}` part would find the whole of 'running', and 'running' is *not* followed by 'ing'. Remember, `{1,}` means 'find as many as possible'.