Give English-language description of the languages generated by the following regular expressions.

Discussion: Break the expression down into pieces, describe the language of each piece, and then build up from there. Methodically writing out strings that match the pattern can also be a useful tactic.

(a) $(a|b)^*$ — any number of as and bs

Discussion: $(a|b)$ is a single a or b, so $(a|b)^*$ means zero or more occurrences of a single a or b.

(b) $a^*|b^*$ — either all as or all bs

Discussion: a^* is any number of as (including 0) and b^* is any number of bs (including 0).

(c) $b^*(ab^*ab^*)^*$ — even number of as.

Discussion: ab^*ab^* is an a followed by any number of bs (including 0), twice — which means a string starting with a and containing exactly two as total. $(ab^*ab^*)^*$ repeats that, for strings starting with a and containing an even number of as (and ϵ). b^* then removes the "starting with a" requirement without changing the total number of as.

(d) b^*abb^* — ab with any number of bs before and after.

Give regular expressions over $\Sigma = \{a, b\}$ that generate the following languages.

(a) $\{x \in \Sigma \mid x \text{ contains 3 consecutive } a\text{'s}\}$ — $(a|b)^*aaa(a|b)^*$

Discussion: aaa is required, so start with a pattern for that (aaa). Then, what can come before or after? Since it wasn't stated that there needed to be exactly 3 consecutive as, it can be any combination of symbols (including none).

(b) $\{x \in \Sigma \mid |x| \text{ is even}\}$ — $((a|b)(a|b))^*$

Discussion: An even number can be written as $2k$ for some integer k, so this provides the idea — a regular expression that matches exactly two characters, repeated 0 or more times.

(c) $\{x \in \Sigma \mid n_b(x) = 2 \bmod 3\}$ — $a^*ba^*ba^*(ba^*ba^*ba^*)^*$

Discussion: This notation means that $n_b(x) = 3k + 2$ for some integer k — that the remainder is 2 when $n_b(x)$ is divided by 3. Proceed similarly to the previous problem — start with a unit that has exactly 2 bs ($a^*ba^*ba^*$), follow it with a unit with exactly 3 bs ($a^*ba^*ba^*ba^*$) and repeat that as many times as desired: $a^*ba^*ba^*(a^*ba^*ba^*ba^*)^*$. It can then be observed that the first a^* in $a^*ba^*ba^*ba^*$ can be dropped — why?

(d) $\{x \in \Sigma \mid x \text{ contains } aaba\}$ — $(a|b)^*aaba(a|b)^*$

Discussion: Start with what is required ($aaba$), then add what can come before and after.

(e) $\{x \in \Sigma \mid n_b(x) < 2\}$ — $a^*|a^*ba^*$

Discussion: Less than 2 bs means there are either no bs or exactly 1 b.

(f) $\{x \in \Sigma \mid x \text{ doesn't end in } aa\}$ — $((a|b)^*(b|ba))|\epsilon|a$

Discussion: Not ending in aa means ending in either b or ba — or ϵ or a.