**Algorithm Explained: Callsign Country Lookup**

Overview

This document describes the callsign lookup algorithm implemented in the get\_cty.py module for the Contest Log Analyzer application. The primary purpose of this algorithm is to determine the correct DXCC entity (country) and associated geographical information (CQ Zone, Continent, etc.) for any given amateur radio callsign.

The entire process is data-driven, relying on standardized "country files" that map callsign prefixes to their respective entities. The system uses two types of these files:

1. cty.dat (The Universal File): This is the standard, comprehensive country file used for all universal annotations. It contains the complete list of DXCC entities, WAE entities, and detailed geographical data. Every log is first processed against this file to establish a consistent baseline of information.
2. Contest-Specific CTY Files (e.g., cqww.cty): Some major contests, like the CQ World Wide DX Contest, publish their own official list of multipliers. These files often have a simplified format and may contain slight differences from the universal cty.dat. When processing a log for a contest that specifies its own country file, the program uses that file to determine the official contest multipliers.

The algorithm follows a strict, multi-step order of operations, stopping and returning an answer as soon as it finds a definitive match or encounters a specific rule that resolves to 'Unknown'.

Step 1: Preparation and Cleanup

Before any lookups happen, the callsign is standardized to ensure consistency.

1. Standardization: The callsign is converted to UPPERCASE, and any characters after a hyphen (-) are removed (e.g., W1AW-1 becomes W1AW).
2. Common Suffix Stripping: A specific list of common, non-location-based suffixes are removed from the end of the callsign. This includes:
   * /P (Portable)
   * /M (Mobile)
   * /B (Beacon)
   * /QRP (Low Power)

*Note: Numeric suffixes like /5 are intentionally not removed at this stage, as they are handled by a special rule later.*

Step 2: High-Priority Overrides

Next, the program checks for a few special cases that take precedence over all other rules.

1. Exact Match Lookup: The highest priority is given to "exact match" callsigns listed in the country file (e.g., =4U1UN for the United Nations HQ). If the cleaned callsign is on this list, its specific entity is returned immediately, and the process ends.
2. Maritime Mobile (/MM): If the callsign ends with /MM, it is identified as a maritime mobile station. According to DXCC and contest rules, these stations do not count as a country, so the result is immediately set to "Unknown".
3. Guantanamo Bay (KG4): Callsigns starting with KG4 are handled by a unique rule. If the callsign is exactly 5 characters long and follows the pattern of KG4 plus two letters (e.g., KG4NA), it is resolved to the specific Guantanamo Bay entity. Any other KG4 callsign is resolved as the continental United States.

Step 3: Slash (/) Handling Logic

If the callsign still contains a slash after the initial cleanup, it enters a detailed set of rules designed for portable operations. The callsign is split into two parts: A (everything before the last slash) and B (everything after the last slash).

A. Special Rule for Numeric Suffixes (e.g., S50A/3)

If B is a single digit, a special set of rules is applied first:

1. US Portable Rule: The program first checks if A has the structure of a standard US callsign (starting with K, W, N, or AA-AL). If it does (e.g., NP2I/4), the location is immediately resolved as the continental United States (primary prefix 'K').
2. Digit Substitution Rule: If A does not have a US structure, the program attempts a digit substitution. It finds the last digit in A (the 0 in S50A) and checks if it's a "call area" digit (i.e., it has letters before and after it). If it is, it replaces that digit with the digit from B to form a new callsign (S53A) and tries to look that up. If successful, that result is returned. If this substitution is not applicable or the resulting lookup fails, the algorithm does not stop; it proceeds to the next set of rules.

B. Standard Slash Rules

If the numeric suffix rule doesn't apply or doesn't yield a result, the standard slash logic begins:

1. Exact Prefix Check: The program checks if either A or B is an *exact* prefix listed in the country file (e.g., EA8 in EA8/G3ABC).
   * If only one of them is an exact prefix, that prefix determines the country, and the process ends.
   * If both are exact prefixes (e.g., EA8/VE3), the location is ambiguous, and the result is "Unknown".

C. Final Portable Logic (CEPT Rule and its Override)

If neither A nor B is an exact prefix, the program enters its final, most nuanced logic:

1. Default CEPT Rule: The default behavior is to follow the CEPT rule, which assumes a format of [Visited Country]/[Home Callsign]. The program looks up the country for A.
2. US/Canada Override: There is one critical exception.
   * If the lookup for A results in a country that is not the US or Canada (e.g., CE3 in CE3/N5NU), the CEPT rule is considered final, and the location is resolved as A (Chile).
   * However, if the lookup for A results in the US or Canada (e.g., KI6RRN in KI6RRN/KL7), the program assumes the operator might be portable in a *different* US or Canadian entity. It then performs a lookup on B (KL7). If B resolves to a valid US or Canadian entity, that more specific location (Alaska) is returned. If not, the original result for A is used. This correctly handles illegal callsigns like W0/EA5JJN by resolving them to the US.

Step 4: Final Fallback - Longest Prefix Match

If a callsign makes it through all of the preceding steps without a match (meaning it has no slashes and isn't a special case), the program performs a standard longest-prefix match. It checks the callsign against the country file database, shortening it by one character at a time until it finds the longest possible matching prefix.