Syntax for ICD-11 postcoordination and clustering

DRAFT 2016-06-14 Vincenzo Della Mea

REQUIREMENTS

The requirement is to establish separators for the syntax used to represent postcoordination and clustering in ICD-11, according to the following basic schemas (3.7.1 in Draft Reference Guide, 2016.03.22), provisionally coded using backslash:

- `stem code`
- `stem code\extension code\extension code`
- `stem code\\stem code`
- `stem code\extension code\\ stem code\extension code\extension code`
- `extension code (history of)\stem code\extension code\extension code\\extension code(family history of)\stem code\extension code`

The key issue is to preserve the above possibilities, using separators that:

- cannot be confounded with codes;
- allow for the double meaning that is currently expressed by using single and double backslash;
- are easy to use (including being known to coders);
- are easy to implement in information systems.

In fact, in information systems and programming languages backslash ("\") is normally used to identify an escape character, and thus may give some issue, although it is possible to use it.

One possible proposal is to use HL7 separator characters: they are well known to health systems developers. The next section will describe them.

HL7 Separator Character Codes

In HL7 messaging, the separator characters are known as the message delimiters or special encoding characters, with a meaning in principle defined as follows (in order of granularity):

|	Field separator ("pipe")
^	Component separator, ("hat")
&	Sub-component separator
~	Field repeat separator
\	Escape character

While these characters are identified as standard for HL7, they are only suggested and even their position in the hierarchy could be changed. In a specific HL7 messaging implementation, the HL7 standard lets you choose your own as long as they are shown at the beginning of the message in the MSH segment.

However, if taken as indicated, two of them could be used for the two levels needed for ICD-11 postcoordination and clustering. If we choose the first two, the examples mentioned at the beginning will become (**A**):

- `stem code^extension code^extension code`
- `stem code|stem code`
- `stem code^extension code|stem code^extension code^extension code`
- `extension code (history of)^stem code^extension code^extension code|extension code(family history of)^stem code^extension code`

Since it is not mandatory to choose these two, other alternatives could be imagined. However, to to easier keyboard placement, the tilde field repeat separator (~) could be avoided.

Thus, for example, a variant that respects the original ordering that seems even more readable is (**B**):

- `stem code&extension code&extension code`
- `stem code|stem code`
- `stem code&extension code|stem code&extension code&extension code`
- `extension code (history of)&stem code&extension code&extension code|extension code(family history of)&stem code&extension code`

Non-HL7 alternatives

If avoiding HL7 suggestions, one possibility, which is particularly readable, is to use the slash character "/" at the first level, and another one (e.g., "&") at the second level, or double slash, although having two different characters could be less confusing (**D**):

- `stem code&extension code&extension code`
- `stem code/stem code`
- `stem code&extension code/stem code&extension code&extension code`
- `extension code (history of)&stem code&extension code&extension code/extension code(family history of)&stem code&extension code`

Precautions

In Iris the pipe is used to indicate a „due to"-relationship between two codes. This has been used in MMDS before as well and therefore has been carried over into the Iris system. It can be changed, but might confuse users of multiple cause data for mortality. To avoid clashes while still using HL7, the suggested alternative could be to use "^" and "&" in any order (depending on how they are perceived by users) (**D**).

When ICD-11 codes have to be transmitted in a HL7 message, if the ICD-11 separators are part of the HL7 syntax too, they have to be escaped according to rules described in section 2.7 of HL7 specifications. In particular: | becomes \F\ , ~ becomes \R\, ^ becomes \S\ & becomes \T\. This will not be made manually, but by software on sender and receiver side.

Recommendation

The most readable solutions are those with “pipe”, “backslash”, and “at” (B,D). Since pipe has already a meaning in the ICD community, solution **D** seems the most adequate, and includes “/” as stem to stem separator and “&” as stem to extension separator.