

Mariposa is talking with some of the others in the computing laboratory.

“So,” says Mariposa “you know how you suggested that I try to write an original poem to go through the language barrier using some of the twenty-six sentences featured in the poetry events well I tried it and I got the following.

It is spring.

It is raining.

It is summer.

It is sunny.

It is autumn.

It is cloudy.

It is winter.

It is snowing.

“So far so good, but then I was thinking about it and I thought - what if it was not snowing but was frosty - I couldn't do that. Now I could ask you to add a sentence for that and maybe you would, but what about thunder, lightning and so on but that is not what I am asking about, it is about the limitations of the localizable sentences method. A sort of analogy is that when using street view on a computer you can walk in a virtual sense around somewhere like Florence and get a good idea of how features which are usually just shown as stand-alone images in places like holiday brochures and so on are all located relative to each other and which shops and whatever are along the route, but ultimately, there comes a place when there is a road that one cannot go along. It seems that it is a similar sort of thing to not being able to write 'It is frosty.' in the poem. Is that right?”

“Yes that seems a good way of putting it.” says John.

“The thing is,” says Caroline “the system does what it does in some particular situations and is very good at what it does in those situations, but does have restrictions in what it can do. One can compare and contrast the system with someone who is not very good at languages with a grammar book but has a comprehensive dictionary. Maybe that person can make themselves understood though the language barrier, even if it is a bit clunky, but they have no provenance that they have said what they meant to say and not something else. Yet with localizable sentences there is provenance within the scope of its supported scenarios, yet only within its specific capabilities. There is no graceful fallback for anything outside the specifically encoded sentences. For example, when I gave my talk in the café about seeking information about relatives and friends, someone asked about in-laws. Since then we have extended the repertoire by a few sentences to cover those cases, but before then it was just not possible to convey that information in a localizable sentences message.”

“Ah, thank you.” says Mariposa. “There is something else I would like to ask about if I may.”

“Please do.” says John.

“Well, I don’t know what it is about but someone was asking if decoding localizable sentences had a sort of ffi problem and someone say ‘no’. I was wondering what that is about please.”

“Ah,” says John, “I think that Henry is the best person here to answer that one. Henry, could you explain please.”

“Ah, yes,” says Henry, “You know how your analogy with street view helped explain about a similar underlying structure sort of situation with localizable sentences, well the ffi problem relates to fontmaking, yet to someone who understands the fontmaking part and knows its well-established solution, reference to it in that way helped to convey the question without needing to go into a lot of detail in asking the question. If I may first explain the fontmaking part please. In advanced typography there is a situation, based on metal typesetting from hundreds of years ago where if there are two or three characters such that their glyphs next to each other in a word can look a bit strange, then sometimes there was a special piece of type that had a combined glyph for the two letters and that was used instead. Nowadays, electronic typography can do the same sort of thing, provided that the font includes a glyph for the ligature as it is called. But it needs to be set up in a table within the electronic font, the glyph substitution table. Now the software using the font searches through the rules that the fontmaker has put into the table and that search is done row by row until it either finds a match to the rule and that tells it which glyph to substitute for the two or three glyphs mentioned in the rule. If it finds a match, then that is it, if there is no match by the time that the search gets to the end of the table, then there is no glyph substitution. Do you follow.”

“Yes.” says Mariposa.

Henry continues “Now the thing is, there are often five particular sequences that each start with a lowercase f that are substituted by a ligature. I’ll write them down.”

ff fi fl ffi ffl

“Ah, thank you.” says Mariposa.

“With the ligatures, typically the i in an fi ligature and the i in an ffi ligature lose the dot.” continues Henry, “But the thing is, if those sequences are listed in that order in the glyph substitution table, then the ffi and the ffl are never substituted, because a search for a match is made and once the rule for ff is found in the table, the glyph for ff is substituted so instead of a ligature glyph for ffi one would get a ligature glyph for ff followed by a letter i on its own. So the ffi and ffl rules need to be before the ff rule and before the fi and fl rules as well. So an order that will get the correct result is as follows.”

ffi ffl ff fi fl

“Ah yes, I understand.” says Mariposa.

“So,” continues Henry, “is how does that relate to localizable sentences. Well, there is a decoding table of code numbers and localized text that is searched sequentially. So, for example, in the scenario about seeking information about relatives and friends after a disaster, one particular code number for a sentence, and its localized version in English is as follows.

313592 The person is female.

“Now, the search method for searching through the sentence.dat file may well be very different from how the search method regarding glyph substitution in a font takes place, so the question would not have any foundation in such a situation, yet to help explain something interesting, even important, about the encoding format, let us hypothetically consider what would happen if the search method were the same as for the search method regarding glyph substitution in a font.

“So the question is effectively asking the following question in relation to that particular code number.

“Are any localizable sentences encoded with any of the code numbers 3 31 313 3135 31359?”

“The answer is no. This is due to the design of the encoding space. In addition there are not, and by design cannot be, any codes longer than six digits that start with 313592.”

“Ah, thank you.” says Mariposa.