The Curse of Representation by Specificity, or: Why There Should Not Be a Transgender Flag Emoji

Author: Charlotte Buff
Mail: irgendeinbenutzername@gmail.com
Submitted: 2019-03-25

1. Background

The publication of document L2/19-080 ("Proposal for Transgender Flag Emoji") marks the first formal acknowledgement of the long-running campaign for the inclusion of a transgender pride flag emoji by the UTC. Speaking as a transgender person, as a vexillologist, and as someone who has been writing to the UTC about adding various gender-related emoji for almost three years, I want to strongly advise **against** implementing that proposal under any circumstances. I also want to use this opportunity to talk about some related topics that may be useful for the emoji encoding process in the future.

I apologise in advance for the verbosity. This may seem like unnecessary tangents at first, but it all ties into the main thesis at the end.

2. History and Context

My very first submission to the Unicode Technical Committee was a comment on their then recent plans to enshrine gender as a fundamental property of human-form emoji in the standard. I made the point that the binary gender model employed by the UTC was not only inadequate for representing humanity, but that recognizing any gender at all would be a huge misstep.

Some time later, I submitted another document, this time urging the UTC to at the very least add *all* necessary gender variants if they couldn't be convinced to eschew gender altogether just so nobody would be erased. This was a move of desperation, because even "full" gender coverage wouldn't have solved the fundamental problem of having gender explicitly marked in the standard.

Both of my comments were dismissed and a few months later, Unicode Emoji 4.0 was released with hardly any acknowledgement of the existence of non-binary people and many sequences needed to represent binary identities absent. In short: The standard was now broken and getting rid of gender altogether off the table, which is why since then I have submitted additional documents always asking for the same thing: For the missing emoji to finally be added.

It is a commonly repeated piece of advice that there are only two valid ways to ask for someone's gender on a form:

1. Unrestricted text field, or
2. Don't.

If we accept the conjecture that emoji need to represent real people, this means there are only two ways gender could have been handled: Either every possible gender needs to be separately encoded, or gender isn't a part of the puzzle at all. The former option is impossible because the set of all

genders is potentially infinite, and the latter option was rejected by the Unicode Technical Committee in 2016.

What happened instead is that a third option was taken: There *is* a finite number of genders to choose from, but male and female aren't the only possibilities. Everything non-binary simply gets lumped together into an "other" category.

The "Male/Female/Other" trichotomy is a terrible approximation of human gender, but it is widely used in a multitude of contexts and many non-binary people are at the very least somewhat okay with it. It *does* cover everything, but in the worst way possible.

Regardless, because this 3-gender model is what the Unicode Standard employs for emoji, the total number of gendered emoji is thus finite. In fact, there are only 42 variants that still need to be encoded at the time of writing. Once these 42 have been added, I will no longer have any reason to contact the UTC about gender. There is a clear problem with a realistic and obvious solution.

I wanted to bring up this example as a contrast to other superficially similar campaigns to make Unicode "more diverse".

3. Diverse Emoji

Gender is not a discrete property, but as we have seen there exist conventions to make it discrete without technically excluding anyone. The same can be said of skin colour, which is obviously a spectrum but has been abstracted in Unicode through the use of the Fitzpatrick scale. Leaving aside how appropriate this mechanism really is, there now exists a clearly defined, unambiguous question to determine whether any given emoji is racially inclusive: Does it support all six Fitzpatrick types?

It does not work like that for two other kinds of Unicode modifiers: Hair style and disability.

Ignoring unique anomalies like Person with Blond Hair, Unicode offers five different types of hair styles: Red, white, curly, bald, and "none". Unlike gender, hair is something visual, and unlike skin tone, it is not a one-dimensional spectrum. Unicode's hair model is not an approximation of human hair variety; it is a small, unprecedented, arbitrary selection of more or less common hair types. Personal IDs don't record hair colour as "Red/White/Other" for instance.

And because hair style now *is* an explicit aspect of human-form emoji, this automatically makes everything that doesn't have its own character or sequence non-existent, just like a purely binary gender model would erase all non-binary people. The otherwise reasonable argument that the unmarked base emoji don't technically specify hair (*"The standard doesn't say that Woman isn't wearing a ponytail."*) can no longer be applied, because the hair components demonstrate that the only way for something to be represented in emoji is by explicitly encoding it. Instead of making emoji more inclusive, Unicode now *excludes* everyone who doesn't have one of four idiosyncratic hair styles.

What is the solution here? There is none. Hair style is infinite, but there is no easy way to put it neatly into boxes that would satisfy user expectations. For every type of hair that may be added in the future, two others still won't be part of the set. I could make the same argument even more potently about disabilities, which are currently being represented by only three modifier-like characters (Probing Cane, Motorized Wheelchair, and Manual Wheelchair) to the detriment of everyone these do not apply to. Even as these emoji were originally being announced, half of the

comments on Twitter were by people complaining that *their* specific disabilities (many of which don't even have any visual presentation in the first place) weren't considered as well.

This modifier approach also has the interesting side effect that the "default human being" is now implied to be perfectly able-bodied at all times because every disability is a dependent modifier, which perhaps was not the intended message of these additions.

The problem with talking about diversity and representation in the context of emoji is that the topic is always brought up with complete disregard for the inherent limitations of the medium we're actually dealing with. The discourse is largely dominated by people who don't understand what emoji are or how they work, yet still demand they behave in exactly the way they envision.

You can make a movie or TV show more diverse because actors are real humans. Emoji are not real humans or even abstractions of real humans; they are abstractions of written symbols: Small, discrete, immutable blobs of text data. A song or poem can be written about arbitrary topics just by combining words on the fly, but every single minute detail one wants to represent in emoji form needs to be individually and independently conceptualised, evaluated, encoded, documented, designed, drawn, sorted, outfitted with keywords and categories, and then somehow included on keyboards in a way that is quick and intuitive to use, ideally for the entire world population.

In the realm of pop culture and entertainment, we accept that a TV show for instance cannot have five hundred separate main characters that all represent a different aspect of the human condition. Every program can't accommodate a cast that consists of at least one character each from every nationality on planet Earth because narratives and relationships would be utterly incomprehensible to anyone watching with that many people running around all the time. That is why there exists more than one TV show, incidentally. But there is only one emoji set – it is called **Uni**code for a reason – so consumers insist that every single defining characteristic humans can possess be present simultaneously and concurrently. And since humans are complex, every one of these characteristics needs to be freely combinable with every other, because real people obviously tend to have more than one attribute at once.

Audiences know that a fictional character who is, say, disabled and queer is representation for both the disabled and the queer communities (if written well that is). You don't get people complaining about a lack of queer representation just because all the queer characters in a show happen to also be disabled. But that is not how emoji are viewed.

An emoji that depicts a black, red-haired woman in a wheelchair is not seen as representative of black people, women, gingers, or wheelchair users; it is seen as representative of that specific combination of attributes and nothing else. If you're a *Latina*, red-haired woman in a wheelchair, this emoji is not for you. Go away and use something else. The consensus among the general public appears to be that an emoji is not representative of me unless it equals myself in every possible aspect precisely.

People want emoji to behave like images because that is how they perceive them on their phones. They want them to be literal selfies they can embed in their messages whenever, wherever. But emoji are not images. In fact, they are the exact opposite of images: Plain text. If one wants Unicode to represent everything anyone cares about, one first has to erase everything that Unicode is and does. You, in essence, have to destroy the very concept of digital text just to achieve

something that other well-established and widely supported standards have been doing infinitely better for decades.

I'm not saying that emoji should not be diverse. I'm saying that emoji literally **cannot** be diverse and any attempt at diversity is inevitably doomed to failure. If we lived in a world of magic where I could simply press a button to make computers do anything I want at no cost, I would most likely consider turning emoji into perfect representations of whatever, but here in the real world we have to accept that some things are either completely impossible, or so unimaginably impractical that the negative fallout would far overshadow any benefits you might reap as a result.

Which brings us to flags.

4. The Case Against the Trans Flag Emoji

Unicode Emoji has two well-defined mechanisms for encoding flags: Regional Indicator Symbols and Emoji Tag Sequences. These are based on region codes from ISO 3166 which form an extensive, but distinctly finite set. For every potential flag emoji, the decision on whether it can be in Unicode hinges on only one unambiguous question: Does it have a region code?

At least, that is the theory. In reality, however, most people are not aware of this, and those who are frequently do not care. All people see is an "image" on their phones, and how hard can it be to create another "image"? Why does Apple (for it is always Apple) think that the Pitcairn Islands are more important than *<thing I care about>*?

This erroneous line of thinking leads to constant requests to the UTC and vendors for all sorts of flags representing geographical entities, ethnic groups, social movements etc. which do not correspond to any region code, be it NATO or Australian Aborigines or Pan-Africanism or Kurdistan or the Esperanto language or – in this case – transgender people.

I oppose the addition of any non-regional flag emoji for one simple reason:

<div align="center">

It will never end.

</div>

The UTC added the rainbow flag as a ZWJ sequence, so now people are (perhaps rightfully even) asking why there isn't a transgender pride flag as well. After all, the only way for something to be represented in emoji is by explicitly encoding it, and while the rainbow flag does serve as an umbrella symbol for the entire LGBT community, there are separate flags for its many subgroups as well. People don't want to be represented as "...and the rest".

So the UTC decides to add the trans pride flag, but people are still complaining where the bi pride flag is.

So the UTC decides to add the bi pride flag, but people are still complaining where the ace pride flag is.

So the UTC decides to add the ace pride flag, but now the UTC is wondering whether there are any pride flags they have missed. (There are, and they will notice because of the continuing complaints.)

Pansexuality and bisexuality are essentially the same thing but have widely different flags. I would know because I am pansexual. But there is only one symbol in the Unicode Standard that signifies "attracted to more than one gender" to be used as a component in ZWJ sequences and it was already

used for the bi pride flag, so that's a problem. Lesbians have their own flag, and it is apparently so obscure that even people advocating for more pride flag emoji tend to forget it exists. What about BDSM culture? That has its own flag – many flags in fact – but it's not really inherently part of queer identity. Still, people will want those. Do furries have a flag, too?

And while this discussion has been going on, the NATO guys, the Aborigine activists, and the Esperanto speakers have been knocking on the UTC's door ranting about their symbols and why Unicode hates them so much.

There is no definitive list of identity flags; the set is ever growing, ever changing, and inherently ill-defined. But if the UTC stops adding more emoji at any point during this process, groups who haven't been included yet will read this as a targeted attack against them in particular. Nevermind the sports clubs and political parties from all over the world that have their own symbols, or the dozens of separatist movements that want to have their flags in Unicode. Every day that passes without the UTC announcing the flag of Somaliland for the next emoji update is a political statement; no transgender flag in Unicode means that the Consortium does not care about transgender people.

No matter how quickly new additions are being released, people will get mad unless you manage to add every single request in one go. But you can't, because the set of all possible "things someone cares about" is infinite. **The only way forward at this point is to make decisions about which types of persons are important enough to be part of what some have called a "universal language" and who needs to be permanently left out.** There is just no way to make that sentence sound nice. There is no success state in this project. Once again, the push to make emoji more inclusive has directly led to emoji being forced to deliberately and inevitably exclude many groups of people.

Adding the transgender pride flag is not going to solve anything. It will only further legitimise the existence of the rainbow flag emoji, thus giving fuel to the idea that everyone and everything needs to be discretely encoded as part of the Unicode Standard, which is impossible to achieve. People are angry now and they will continue being angry because they do not realise that this is a problem without a solution.

5. Conclusion

As I have explained, I am not arguing against the concept of diverse representation. Under any other circumstances, this document would only consist of the words "Yes, please". I am, however, arguing against diverse representation in places where it physically cannot exist, and emoji is one of them. Perhaps it is even the only medium where diversity as we understand it cannot ever happen.

Humans can represent other humans. Facsimiles of humans can represent other humans. You can take a photo of every one of the 7.6 billion people alive today and you would only have exhausted about $10^{-10^{7.8}}\%$ of all possible pixel configurations in a tiny 72×72 PNG image (the preferred format for emoji submissions). But emoji are not images, and we can't have ZWJ sequences that are hundreds of codepoints long.

It does not matter how many concepts are being added to Unicode as new emoji because the number of concepts that *aren't* included will never meaningfully decrease.

The only way for emoji to be truly inclusive and representative is for them to be as generic and nondescript as possible. If even one person can point at an emoji and say "This one looks just like me!", then the designers have already failed. Human-form emoji should not have any skin colour, any gender, or any hair, they should be neither abled nor disabled. I know that is not what the people want, but people tend to want a lot of things that are ultimately bad for everyone involved.

Folks don't even use these emoji to communicate anymore. Modern emoji exist for the sole purpose of decorating one's username on Twitter. They have become posters on a teenager's bedroom walls. This is not a good use of the UTC's or anyone's time and resources.

Every decision to include someone or something in the emoji set is inevitably a decision to exclude everyone or everything else. Always. The more emoji are being added for the sake of perceived diversity, the worse this problem becomes, for there is never a last emoji. You cannot win at this; nobody wins at this.

This is why I wrote that document all the way back in 2016. This is what I have been trying to warn you about all this time. Once you set the precedent that a specific group of people can be added to this set of pictographs for no other reason than to have them there, *every* group of people wants the same piece of the cake. And they will judge you if they don't get it, no matter how good your reasons may be.

You added skin tone modifiers because vendors made terrible fonts, but as a consequence people now want every single defining characteristic of human beings to be modifiers as well because they have seen that something like this is apparently possible; gender, hair, and disability are only the tip of the iceberg. You added one solitary non-regional flag as a ZWJ sequence because people seemed interested in it, but now every single person who has ever rallied under any sort of flag wants their insignia in the standard as well because they have seen that it works.

This is what I call The Curse of Representation by Specificity. Through your actions you have irrevocably declared that only those who are explicitly mentioned as their own distinct entity in the standard are considered worthy, so now the only way for you to not be exclusionary towards any particular group is by becoming ever more specific and disunifying said group from all the rest, even though your framework does not remotely support this approach by any reasonable means. The moment you say no to anyone for any reason, you will be seen as the villain, thus you are spellbound to simply add more and more stuff that fewer and fewer people will actually use until everything inevitably collapses under the load.

So please, as a transgender person who has been very pushy about getting additional gendered emoji added, do not entertain the idea of accepting the transgender flag emoji. Too much damage has already been done; better to stop this now before it gets even more out of hand.

Because only if nobody is represented, everybody is.