

Source:	CheonHyeong Sim (沈天珩, aka TianHeng Shen)
Title:	Response to IRGN2735
Status:	Individual Contribution on IRG #63
Action:	To be considered by IRG and multiple sources

Background

In IRGN2735, Wang Xieyang (王谢杨) expressed some opposing opinions on disunifying U+2335F (冑) and U+6B25 (𠂔). In this document, I would like to respond to that feedback item by item, and before that, I would like to reaffirm that Unicode encodes *characters* but not *glyphs*.

Responses

I think there is no **urgency** or necessity to disunify the two characters at this time.

...

In fact, users have already widely used character U+6B25 to display 𠂔日欠 and 𠂔日欠. This action **has not caused any conflict or inconvenience**.

About the “necessity”, I would like to respond later in this proposal; however, about the “urgency”, I would like to say that in fact most of the disunified characters are not that urgent. Do you think that 𩇛 (U+9FEA) is very urgent to be disunified from 𩇛 (U+3E02)? Do you think that 𠂔/𠂔/𠂔/𠂔 (U+4DBA..U+4DBD) are very urgent to be disunified from 𠂔/𠂔/𠂔/𠂔 (U+80AD / U+43D9 / U+440B / U+6721)? Do you think that 𠂔 (U+2B738) is very urgent to be disunified from 𠂔 (U+53F1)? These are all not common characters, their usages are very limited, and the past unification also “did not cause any conflict or inconvenience”. To sum up, you *should not* oppose the disunification by their *non-urgency*.

冑，舊註“音猥，吐也”。按，《說文》：“𩇛，歐貌。心惡未至于歐，因𩇛出之。”或省作「𠂔士同」。舊註義與「𠂔士同」近，音與𩇛別，故知其譌。又與肉部昌溷。𠂔非。——《正字通》

The quoted text from 正字通 says that this character is **derived from the left component of 𩇛 (U+55C0)**. The explanation is reasonable.

猥 影蟹賄一合 上聲10賄
設 曉通屋一開 入聲1屋
曉江覺二開 入聲3覺
言 影臻質A三開 入聲4質

I have doubts about whether 冃 (月部) is really derived from 設 or 言.

The radical of 𠄎 (U+6B25) in the code chart now is lack (欠), so the change of its left component will not change its radical.

Meanwhile, it is easy to notice that there is still no V source reference for this character.

...

I suggest that IRG maintain the current situation and recommend Vietnam, if needed, propose a horizontal extension of 𠄎日欠 to 𠄎 (U+6B25) while **having lack (欠) as its radical**.

The radical of 𠄎 (U+20BD0, CJK-ExtB) and the radical of 𠄎 (U+31488, CJK-ExtH, [ORT link](#)) are both 口, so unify 𠄎 to 𠄎 will not change its radical. Meanwhile, 𠄎 has only G-, T- and J-source, and 𠄎 has only V-source. Obviously it would be more convenient for Vietnam to horizontally extend 𠄎 to 𠄎 if you do not consider their non-cognition academically, however, the fact is that they were not unified.

What is more, we always use the semantic component (形旁) as the radical and never use the phonetic component (声旁) for all the phonograms (形声字). It is no doubt that the semantic component of 𠄎日欠 (either the one from 康熙字典/大漢和辭典 or the another from Nôm) is 日, so we have no reason to “have 欠 as its radical”.

Again, Unicode encodes characters but not glyphs, so once we could determine they are different characters, no matter how similar their glyphs are, they should be separated. I also said this sentence to Wang Xieyang a month ago when he was suggesting to unify 𠄎口、 (WS2024-00815, [ORT link](#)) to 𠄎口、 (U+20B9A). For 𠄎口、, not only I, but also Andrew West and L. F. Cheng oppose the unification, and finally Wang Xieyang was persuaded by us. I believe that 𠄎日欠 and 𠄎日欠 is a very similar case.

Although the semantic component of 「𠄎日欠」 as Chữ Nôm is 日 (U+65E5), the VNPF has already chosen to use 𠄎 (U+6B25) to represent 「𠄎日欠」.

That does not matter. Before the disunification for U+2B735, nomfoundation.org along with the other sites using Chữ Nôm all choose to use 𠄎 (U+722B) to represent “làm” (means “do” in Vietnamese). However, at least, nomfoundation.org has already revised to use the new codepoint to represent it after the disunification. Nothing inconvenient.

Meanwhile, in other books, 𠄎 (U+6B25) is also used as a variant of 吹 (U+5439), and the left component is **clearly** 日 (U+65E5).

You have confused the concepts of *character* and *glyph* again. The semantic component for it as a variant of 吹 is **clearly not** 日. There is no connection between “blow” and “sun”. Considering that an error form does not have an abstract shape (aka “讹字无抽构”, Kushim Jiang, 2022), this example does not illustrate anything.

(End of document)