

Confusables in Balinese, Javanese, and Kawi

Norbert Lindenberg

2023-01-14

Proposal

This document proposes to add a number of pairs of confusable character sequences for the Balinese, Javanese, and Kawi scripts to the [data files](#) for [Unicode Technical Standard #39 : Unicode Security Mechanisms](#), and to remove several entries from the list of intentional confusable mappings.

The following changes are proposed:

- Entries in tables “no longer intentional confusable mappings” should be removed from the file intentional.txt, but kept in the file confusables.txt. Type designers do not always use the same glyphs for these pairs.
- Entries in tables “new intentional confusable mappings” should be added to both the file intentional.txt and the file confusables.txt. Pairs in light gray are already registered.
- Entries in tables “new confusables” should be added to the file confusables.txt, but not to the file intentional.txt. Pairs in light gray are already registered.

Some of the pairs involve character sequences of length greater than 1 as both source and target; such character sequences have also been proposed in [L2/22-107](#) and [L2/22-108](#), identifying the need to support such pairs in UTS #39.

This document uses the fonts Noto Sans Balinese and Noto Serif Balinese for Balinese, Noto Sans Javanese and Tuladha Jejeg for Javanese, and Tantular Kawi for Kawi.

Balinese – no longer intentional confusable mappings

Source	Rendering	Target	Rendering	Comments
1B28	ᮘ / ᮙ	1B58	ᮘ / ᮙ	Source and target differ in Noto Sans Balinese, but use the same glyph in other fonts.
1B50	o	1B5C	o	May or may not differ in size.

Balinese – new intentional confusable mappings

Source	Rendering	Target	Rendering	Comments
1B3C	ꦱꦶꦒ	1B44 1B2E 1B42	ꦱꦶꦒ	The lower part of SIGN LA LENGA is the conjunct form of LETTER LA; the upper part is SIGN PEPET.
1B0D	ꦥꦺꦥꦺꦠ	1B52	ꦥꦺꦥꦺꦠ	
1B11	ꦥꦺꦥꦺꦠ	1B53	ꦥꦺꦥꦺꦠ	
1B0F	ꦥ	1B56	ꦥ	
1B61	ꦲ	1B35	ꦲꦲ	Several Balinese musical symbols reuse the shapes of combining marks. Confusion is unlikely when there's a contrast between spacing and nonspacing glyphs, but likely where both are spacing.
1B62	ꦲ	1B3E	ꦲꦲ	
1B6A	ꦲ	1B04	ꦲꦲ	

Balinese – new confusables

Source	Rendering	Target	Rendering	Comments
1B1B	𑌎	1B10	𑌎	The small difference in the upper right terminal may or may not be present; U+1B1B is very rare.
1B5F	𑌕	1B5E 1B5E	𑌕𑌕	Differ only in spacing.
1B60	𑌖	1B54	𑌖 / 𑌖	U+1B54 may or may not have left-side tail.
1B75	𑌜	1B50	𑌜	May or may not differ in size.

Javanese – no longer intentional confusable mappings

Source	Rendering	Target	Rendering	Comments
A9D0	0	A9C6	0	May or may not differ in size.

Javanese – new intentional confusable mappings

Source	Rendering	Target	Rendering	Comments
A9D1	ᮊ	A992	ᮊ	
A98A	ᮊᮒ	A994 A9C0 A9A4	ᮊᮒ	The upper part of LETTER NGA LELET is LETTER NGA; the lower part is the conjunct form of LETTER NA.
A9D6	ᮊᮓ	A98C	ᮊᮓ	
A9D7	ᮊᮔ	A9AD	ᮊᮔ	
A9D9	ᮊᮕ	A9AA	ᮊᮕ	

Javanese – new confusables

Source	Rendering	Target	Rendering	Comments
A98A	ᮊᮒ / ᮊᮓ	A9D2	ᮊᮒ / ᮊᮓ	May or may not differ in below-base part.
A9A3	ᮊᮔ	A99D	ᮊᮔ	
A9C9	ᮊᮕ	A9C8 A9C8	ᮊᮕ	Differ only in spacing.
A9CB	ᮊᮖ	A9CA A9CA	ᮊᮖ	Differ only in spacing.
A9C0 A9B0	ᮊᮗ	A9C0 A9B1	ᮊᮗ	Glyphs are swapped between standard Sriwedari and traditional Mardi Kawi orthographies.
A9C0 A9A8	ᮊᮘ	A9C0 A9A7	ᮊᮘ	Glyph for conjunct form of U+A9A7 in standard Sriwedari orthography is used for conjunct form of U+A9A8 in traditional Mardi Kawi orthography.
A9D3	ᮊᮙ / ᮊᮚ	A994 A9BE	ᮊᮙ / ᮊᮚ	The upper part of DIGIT THREE is LETTER NGA; the lower-right part may be PENGKAL.
A9D8	ᮊᮛ / ᮊᮜ	A9A6	ᮊᮛ / ᮊᮜ	May or may not differ in tail.

Kawi – new confusables

Source	Rendering	Target	Rendering	Comments
11F44	ᳵ	11F43 11F43	ᳵ	Differ only in spacing.
11F48	ᳶ	11F54	ᳶ	May or may not differ, depending on style.
11F4A	᳷	11F03	᳷	
11F4C	᳸	11F50	᳸	May or may not differ in size.
11F4E	᳹	11F51	᳹	

Cross-script – new confusables

Source	Rendering	Target	Rendering	Comments
11F50	᳸	1B50	0	May or may not differ in size.
A9D0	᳸	1B50	0	May or may not differ in size.
A9D4	ᳶ	1B54	3 / ᳶ	U+1B54 may or may not have left-side tail.
A9D5	᳷	1B55	᳷	