

コーパス内の文字出現頻度による 言語間漢字変換ツールの作成

Agro Rachmatullah 小川泰弘 外山勝彦
(名古屋大学)

中国 簡体字	台湾 繁体字	香港 繁体字	韓国 繁体字		日本 新字体
骨 U+9AA8	骨 U+9AA8	骨 U+9AA8	骨 U+9AA8		骨 U+9AA8
号 U+53F7	號 U+865F	號 U+865F	號 U+865F		号 U+53F7
证 U+8BC1	證 U+8B49	證 U+8B49	證 U+8B49		証 U+8A3C
恶 U+6076	惡 U+60E1	惡 U+60E1	惡(악) U+60E1	惡(오) U+F9B9	悪 U+60AA
卫 U+536B	衛 U+885B	衛 U+885E	衛 U+885B		衛 U+885B
叙 U+53D9	敘 U+6558	敘 U+654D	敘 U+654D		叙 U+53D9
乡 U+4E61	鄉 U+9109	鄉 U+9109	鄉 U+9115		郷 U+90F7

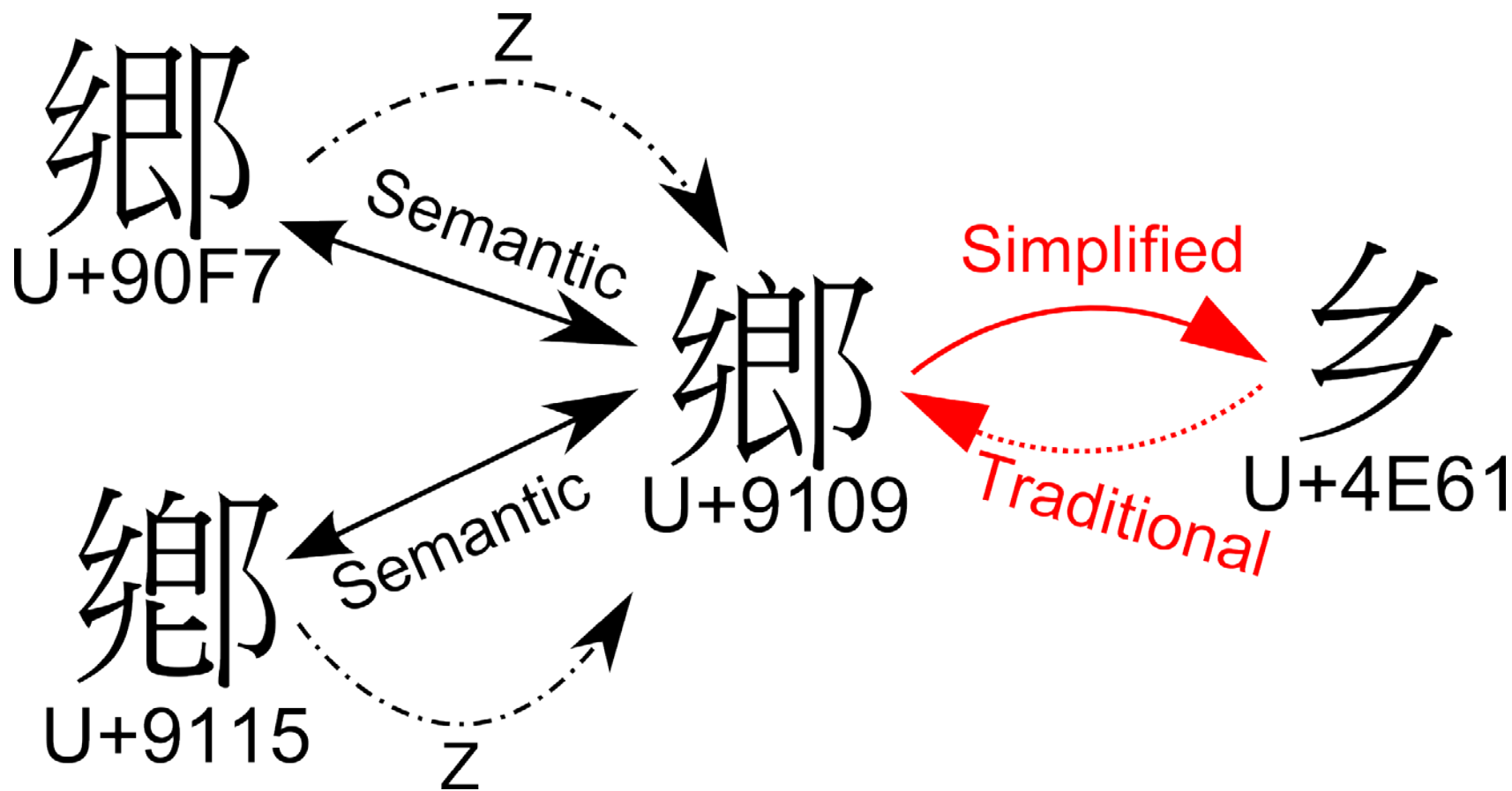
中国 簡体字	台湾 繁体字	香港 繁体字	韓国 繁体字		日本 新字体
骨 U+9AA8	骨 U+9AA8	骨 U+9AA8	骨 U+9AA8		骨 U+9AA8
号 U+53F7	號 U+865F	號 U+865F	號 U+865F		号 U+53F7
证 U+8BC1	證 U+8B49	證 U+8B49	證 U+8B49		証 U+8A3C
恶 U+6076	惡 U+60E1	惡 U+60E1	惡(악) U+60E1	惡(오) U+F9B9	悪 U+60AA
卫 U+536B	衛 U+885B	衛 U+885E	衛 U+885B		衛 U+885B
叙 U+53D9	敘 U+6558	敘 U+654D	敘 U+654D		叙 U+53D9
乡 U+4E61	鄉 U+9109	鄉 U+9109	鄉 U+9115		郷 U+90F7

中国 簡体字	台湾 繁体字	香港 繁体字	韓国 繁体字	日本 新字体
骨 U+9AA8	骨 U+9AA8	骨 U+9AA8	骨 U+9AA8	骨 U+9AA8
号 U+53F7	號 U+865F	號 U+865F	號 U+865F	号 U+53F7
证	證	證	證	証

5つの地域に渡る字体変換ツールを作ること

U+6076	U+60E1	U+60E1	U+60E1	U+F9B9	U+60AA
卫 U+536B	衛 U+885B	衛 U+885E	衛 U+885B	衛 U+885B	衛 U+885B
叙 U+53D9	敘 U+6558	敘 U+654D	敘 U+654D	敘 U+654D	叙 U+53D9
乡 U+4E61	鄉 U+9109	鄉 U+9109	鄉 U+9115	鄉 U+9115	郷 U+90F7

アプローチ



アプローチ

