

Universal Multiple-Octet Coded Character Set International Organization for Standardization

Doc Type: ISO/IEC JTC1/SC2/WG2/IRG
 Title: Proposal to modify source references of V-Source ideographs
 Source: Vietnam
 Status: National Body Proposal
 Authors: Eiso Chan, Lee Collins, Ngô Trung Việt
 Action: For consideration by the IRG
 Date: 2021-04-26

Vietnam proposes modifications to the conventions for IRG source references for V-Source ideographs. The new scheme would be the following:

1. Retain the existing V0 to V4 prefixes.
2. Replace all source references currently having the V or VU prefix with a new VN prefix, which is followed by a 5-digit hexadecimal code. The value of the 5-digit hexadecimal code would be the Unicode code point, including those in the SPUA. Codes in planes 0, 2 and 3 would be primarily from the common set of ideographs used throughout East Asia and added to the URO or extensions through horizontal extensions. Codes in the SPUA would be those originally identified for use primarily in Vietnam, especially not exclusively in the Nôm script, and assigned pending encoding in the URO or extensions. The SPUA codes would be retained permanently, even after extension to the URO or extensions.

The attached *kIRG_VSource_update.txt* data file includes all of the proposed source reference changes per the above conventions. The third column is the current (Unicode Version 14.0 Alpha) property value, and the fourth column is the proposed new property value.

Consequently, we propose the following changes to the documentation of the kIRG_VSource property in the UniHan database (UAX #38):

Property	kIRG_VSource
Status	Normative
Category	IRG Sources
Introduced	3.0
Delimiter	N/A
Syntax	V[0-4N]-[023F]?[0-9A-F]{4}

Description	The IRG “V” source consists of data from the following national standards and other reference works from Vietnam. 1. V0: TCVN 5773:1993 2. V1: TCVN 6056:1995 3. V2: VHN 01:1998 4. V3: VHN 02:1998 5. V4: "Kho chữ Hán Nôm Mã hoá," Hà Nội, 2007. This reference work contains detailed descriptions of primary sources. 6. VN: Additional Vietnamese horizontal and vertical extensions to Unicode not defined in the above sources; the value is its code point. SPUA values primarily represent non-Unicode encoded characters used in Vietnamese digital media and proposed for standardization in Unicode via the IRG working set process, beginning with WS2017. Major sources for VN characters are: a. Nguyễn Quang Hồng. "Tự Điển Chữ Nôm Dẫn Giải." Hà Nội, 2014 b. Anthony Trần Văn Kiệm "Giúp đọc Nôm và Hán Việt," Đà Nẵng, 2004 A stable reference in chart-form is provided here: (provisionally at http://science-technology.vn/)
-------------	---