Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Considering languages in encoding Cuneiform
Source: Michael Everson
Status: Expert Contribution
Date: 2003-06-01

Discussions on the encoding of cuneiform have been ongoing since at least 1999, but with little result. Chief among the reasons for this has been the preoccupation with the glyph representation of cuneiform signs, which – naturally enough – presents quite a problem. But this problem is orthogonal to the question of encoding.

It will be difficult to encode Cuneiform all at once. Cuneiform was used for 3,000 years for 15 different languages. At any given period or place between ca. 300 and 900 signs or more were used, depending on the language depending upon the language or dialect and period. Old Babylonian, for instance, used about 900 signs; Neo-Assyrian about 600. Unification is both desirable and reasonably straightfoward. Most of the signs are commonly used. Taking Akkadian as a base, at least 50% of its signs are used in all languages. The reading of signs may differ greatly depending on location, language, and time: sign 298 𒄭 (in von Soden's catalogue *Das Akkadische Syllabar*) was read *mim*, *rag*, and *sal* in Akkadian but was read *šal* and in Hittite and *šel* in Hurrian (sign 297 in Rüster & Neu's catalogue). Sign 22 𒀭 'city' in Sumerian is *uru*; in early Akkadian this was read *ālum*, and in Hittite (sign 229) it is read *ḫapiru*.

The term *Akkadian* refers to Old Akkadian, Old Assyrian, Middle Assyrian, Neo-Assyrian, Old Babylonian, Middle Babylonian, Neo-Babylonian, and Late Babylonian.

It is proposed here to ignore supplementary readings as well as glyph variants while preparing the basic code table for Cuneiform. Unification of this sort is exactly what we have done for the CJK characters. An example of this is the character 山 'mountain', which in Chinese is pronounced *shān*, but is pronounced *san* or *yama* in Japanese, and *san* in Korean. For Cuneiform, what is proposed is to take 325 Akkadian characters from von Soden's catalogue as a base (on the advice of Dr Petr Vavroušek, Prague), and them compare them with Hittite, Hattian, Hurrian, Palaian, Luvian, Elamite, Old Akkadian, Old Assyrian, Middle Assyrian, Neo-Assyrian, Old Bablyonian, Middle Babylonian, Neo-Babylonian, Late Babylonian, Ugaritic (not the Ugaritic alphabetic script), and, finally, Sumerian.

The table below begins this endeavour. In the first and second columns the von Soden's glyph and catalogue number are given. In the Hittite, Hattian, and Hurrian columns, Rüster and Neu's catalogue numbers are given, together with one (though there may be many) of their readings for the character; this is a Hittite reading where available, otherwise it is Sumerian and given in capital letters as is customary. At the end of the document characters present in Rüster & Neu but not in the von Soden list used here are given, but as of this date I have only added to sign 121;

Glyph	Akkadian	N°	Hit	Hat	Hur	Pal	Luv	Ela	OAk	OAs	MAs	NAs	OBa	MBa	NBa	LBa	Uga	Sum
𒀀	AŠ	001	001 aš															
𒀁	ḪAL	002	002 ḫal															
𒀂	MUG	003	022? MUG															
𒀃	BA	004	205 ba															
𒀄	ZU	005	209.1 zu															
𒀅	SU	006	213 SU															
𒀆	ŠUN	007																

	BAL	008	004 bal															
	GIR2	009	006 GÍR															
	BUR2	010																
	TAR	011	007 tar															
	AN	012	008 an															
	NANNA	013																
	IT2	014																
	KA	015	133 ka															
	DILIB2	016																
	BUM	017																
	BUN	018																
	EME	019	147 EME															
	MU3	020																
	NAG	021	148 NAG															
	URU	022	229 URU															
	URU2	023																
	GIŠGAL	024																
	ARAD	025	016 ARAD															
	ARAD2	025A																
	ŠAḪ	026	309 šaḫ															
	LA	027	095 la															
	APIN	028	009 APIN															
	MAḪ	029	010 maḫ															

	TU	030	346 tu															
	LI	031	343 li															
	PAP	032	256.1 PAB															
	PUŠ2	033	256.2 PAB-ḪAL (PÚŠ)															
	BULUG3	034	257 BÙLUG															
	MU	035	017 mu															
	QA	036	012 qa															
	KAD2	037																
	KAD3	038																
	GIL	039																
	KID2	040																
	RU	041	043 ru															
	BAD	042	013 pát															
	NA	043	015 na															
	ŠIR	044																
	NUMUN	045	012 kul															
	TI	046	037 ti															
	MAŠ	047	020 pár*															
	BAR	048	020 pár															
	NU	049	011 nu															
	MAŠ2	050	038 MÁŠ															
	KUN	051																
	ḪU	052	024 ḫu															

	U5	053	025 U5															
	NAM	054	039.1 nam															
	IG	055	067 IG															
	MUD	056	026 (mut)															
	SA4	057																
	RAD	058	029 rad															
	ZI	059	033 zi															
	GI	060	030 gi															
	RI	061	032 ri															
	NINNI7	062																
	NUN	063	036 NUN															
	KAB	064	049 kab															
	ḪUB2	065	049* kab															
	ḪUB	066	050 ḫub															
	GAD	067	173 gad															
	DIM	068																
	MUN	069	018 MUN															
	AG	070	081 ag															
	EN	071	040 en															
	DAR3	072																
	SUR	073	042 šur															
	SUḪ	074																
	INANNA	075	041 INANNA															

	SA	076		200 sa	200 sa													
	GAN2	077	061 gán															
	KAR2	078																
	TIK	079	201 GUN?															
	DUR	080	202 dur															
	GUN	081	201 GUN?															
	LAL3	082	170 LÀL															
	DAR	083																
	GUR	084	185 gur															
	SI	085	086 ší															
	SU4	086																
	SAG	087	192 šag															
	MA2	088	087 MÁ															
	DIR	089																
	TAB	090	090 tap															
	LIM2	091																
	TAG	092	091 šum															
	AB	093	097 ap															
	NAB	094	100 nab															
	MUL	095	100 MUL															
	UG	096	093 ug															
	AZ	097	092 az															
	URUDU	098	109 URUDU															

	KA2	099	167 KÁ															
	UM	100	098 um															
	DUB	101	099 tup															
	TA	102																
	I	103	217 i															
	IA	104	218 ya															
	KAN	105																
	KAM2	106																
	TUR	107	237 TUR															
	AD	108	105 ad															
	ṢI	109																
	IN	110	354 in															
	RAB	111																
	LUGAL	112	112 LUGAL															
	ŠIR3	113	106 ŠÌR															
	BAD	114																
	SUM	115	350 SUM															
	KAS	116	259 KASKAL															
	GAB	117																
	EDIN	118																
	DAḪ	119																
	AM	120	168 am															
	UZU	121	203 UZU															

	NE	122	169 ne															
	ERIM2	123																
	BIL2	124																
	ŠAM3	125	103 šàm															
	RAM	126																
	ŠAM2	127	123 ŠÁM															
	ZIG	128																
	KUM	129																
	GAZ	130	122 gaz															
	UR2	131	124 úr															
	SUḪUŠ	132																
	KAŠ4	133	129 KAŠ4															
	IL	134																
	DU	135	128 du															
	LAḪ4	136	236 LAḪ4															
	TUM	137																
	GEŠTIN	137A	131 wi5															
	UŠ	138																
	IŠ	139																
	BI	140	153 pí															
	ŠIM	141	154 ŠIM															
	KIB	142																
	NA4	143																

	TAK3	144																
	KAK	145																
	NI	146																
	NI-A	146A																
	IR	147	077 ir															
	MAL	148	056 MAL															
	KISAL	149	248 KISAL															
	UR3	150	058 ÙR															
	PAR3	151																
	DAG	152	243 tág															
	PA	153	174 pa															
	ŠAB	154	175 šab															
	SIPA	155	177 SIPA															
	GIŠ	156	178 eš															
	GIŠ-BAR	157																
	BIL3	158																
	GIŠ-TUKUL	159																
	AL	160	183 al															
	UB	161																
	MAR	162	191 mar															
	E	163	187 e															
	DUG	164	162 DUG															
	UN	165	197 un															

	KID	166	194 KID															
	ŠID	167	231 ŠID															
	MES	168																
	U2	169	195 ú															
	GA	170	159 ga															
	IL2	171	161 ÍL															
	LUḪ	172	198 luḫ															
	KAL	173	176 kal															
	E2	174	199 É															
	NIR	175	204 nir															
	GI4	176	234 GI4															
	GIGI	177																
	RA	178	233 ra															
	DUL3	179																
	LU2	180																
	LU2-BAD	181																
	ŠIŠ	182																
	NANNA2	183																
	SAR	184	353 šar															
	ZAG	185	238 ZAG															
	QAR	186	240 gàr															
	ID	187																
	TI8	187A	215 id															

	LIL	188	127 LIL															
	MURU2	189	110 MÚRU															
	DE2	190																
	DA	191	214 da															
	AŠ2	192	241 tàš															
	MA	193	208 ma															
	GAL	194	242 gal															
	BARA2	195																
	GUG2	196	220 GÚG															
	GIR	197																
	MIR	198																
	BUR	199	245 bur															
	SIG7	200	239 SIG7															
	DUB2	201	130 DÚB															
	ŠA	202	158 ša															
	ŠU	203	068 šu															
	ŠU-II	204																
	KAD4	205																
	KAD5	206																
	LUL	207	351 LUL															
	ŠAG5	208																
	GE23	209																
	GAM	210	247 GAM															

	KUR	211	329 kur															
	ŠE	212	338 še															
	BU	213	339 pu															
	UZ	214	340 uz															
	ŠUD	215	341 SUD															
	MUŠ	216	342 MUŠ															
	TIR	217	344 tir															
	TE	218	249 te															
	KAR	219	250 kar															
	LIŠ	220	286 liš															
	UD	221	316 ud															
	E3	222																
	PI	223	317 wa															
	ŠA3	224	294 ŠÀ															
	UḪ2	225																
	ERIM	226	327 ERIM															
	PIR2	227																
	ZIB	228																
	ḪI	229	335 ḫi															
	ŠAR2	230																
	TI2	231																
	DUG3-GA	232																
	AḪ ??	233																

	AḪ	234	332 aḫ															
	KAM	235	355 kam															
	IM	236	337 im															
	BIR	237	334 BIR															
	ḪAR	238	333 ḫar															
	ḪUŠ	239	348 ḪUŠ															
	SUḪUR	240	349 SUḪUR															
	ZUN	241																
	U	242	261 u															
	UGU	243	272 UGU															
	LID	244																
	KIR6	245																
	KIR2	246																
	KIŠ	247	274 kiš															
	MI	248	267 mi															
	GUL	249	271 gul															
	NA2	250	314 NÁ															
	NIM	251			074 nim													
	TUM3	252	275 TÙM															
	KIR7	253																
	LAM	254	306 lam		306 lib													
	ZUR	255																
	PAN	256	118 PAN															

	GIM	257	165 GIM															
	UL	258																
	GIR3	259	259 GÌR															
	DUGUD	259A	268 DUGUD															
	GIG	260	269 GIG															
	IGI	261	288 ši															
	PAD3	262																
	AR	263	289 ar															
	U3	264	265 Ù															
	ḪUL	265	290 ḫul															
	DI	266	312 di															
	DUL	267	267 DUL															
!	DU6	268	211 DU6															
"	KI	269	313 ki															
#	DIN	270																
$	ŠUL	271																
%	KU3	272	069 KÙ															
	PAD	273	295 PAD															
	MAN	274	296 man															
	EŠ	275																
	DIŠ	276	356 diš															
	LAL	277	358 lal															
	LAL2	278		362 lál														

	UŠUR	279																
	LAGAB	280	179 ḫab															
	ZAR	281	181 zar*															
	ḪU3	282																
	TUL2	283	180 túl															
	BUL	284																
	SUG	285	182 SUG															
	NIGIN	286																
	ME	287	357 me															
	MEŠ	288	360 meš															
	IB	289	044 ib															
	KU	290																
	ŠE3	291	212 ŠÈ															
	LU	292																
	DIB	293	210 lu		210 dib													
	KIN	294	047 KIN															
	ŠIG2	295																
	ŠU2	296	251 šú															
	ḪUL2	297																
	SAL	298	297 šal		297 šel4													
	ZUM	299	300 zum															
	NIN	300	299 NIN															
	DAM	301	298 dam															

	GU	302	304 gu															
	GEM3	303																
	UḪ3	304																
	NIG	305																
	EL	306																
	LUM	307	310 lum															
	SUḪ3	308																
	TUK	309	053 TUK															
	UR	310	051 ur															
	A	311	364 a															
	AM3	312																
	ER2	313																
	ID2	314	365 ÍD															
	A-A	315																
	ZA	316	366 za															
	ḪA	317	367 ḫa															
	GUG	318																
	ḪA-ŠU2 (?)	319																
	SIG	320	255 SIG															
	UR4	321																
	ṬU	322																
	NIG2	323	369 NÍG															
	IA2	324	371 IÁ															

𒌋𒌋𒌋	AŠ3	325	372 ÀŠ															
			003 PÉŠ															
			005 šir															
			014 dim															
			019 NAR															
			023 UZ6															
			027 SÈD															
			028 SIxSÁ															
			033 zi															
			034 TÙR															
			035 KUN															
					ge-e													
			39.2 BURU5															
			045 U8															
			046 zul															
			048 šušana															
			052 GIDIM															
			054 SILA4															
			055 BÚGIN															
			057 AMA															
			059 ÀRAḪ															
			060 GALGA															
			062 ERIN															

			063 ŠÉŠ															
			064 ŠU.NÍGIN															
			065 SÍG															
			066 SÍG-MUNUS															
			070 GIŠIMMAR															
			071 DÀRA															
			072 ni															
			073 IA4															
			075 GAG															
					076 né-e													
			078 LÚ															
			079 SES															
			080 AŠGAB															
			082 MÈ															
			083 GÙN															
			084 ITI															
			085 ŠINIG															
			088 ŠÙDUL															
			089 DIRI															
			094 NIB															
			096 UKU															
			102 DÉ															
			104 AZU															

			107 EZEN4															
			108 zé															
			111 UNU															
			112 miš															
			113 ḫe															
			114 BÀD															
			116 DÌM															
			117 il															
			119 NINDA															
			120 kum															
			121 ÁG															