

Archaic cuneiform numbers

Robin Leroy, Anshuman Pandey, and Steve Tinney

2024-08-05

Contents

1	Summary	2
2	Proposed changes to the Standard	2
2.1	Summary of proposed characters	2
2.2	Properties	2
2.3	Character names list	2
2.4	Core specification text	2
3	Rationale for curviform–cuneiform disunification	2
3.1	The cuneiform encoding model	3
3.2	Arguments for curviform–cuneiform unification	3
3.3	Metrology	4
3.3.1	The discrete counting system	5
3.3.2	The area system	5
3.3.3	The capacity system	6
3.3.4	The length system	7
3.3.5	Fractions	7
3.4	Early metrology	7
3.4.1	Use in modern publications	9
3.5	Non-numeric usage	9
3.6	Limited benefits of diachronic encoding for numerals	9
3.6.1	Compatibility with transliteration	9
3.7	Compatibility considerations	9
3.7.1	The case of ŠAR ₂	9
4	Rationale for ED–Uruk numeral unification	9
5	Considerations on individual numeral series	9
6	Characters not included in this proposal	9
6.1	Missing numerals	9
6.2	Stacking patterns	9
6.3	Matters for higher-level protocols	10
7	Acknowledgements	10

	Uruk III & earlier	ED – Ur III	OB & later
Numerals	This proposal		
Non-numeric signs	Future Pcun	Existing Xsux	

Table 1: Usage of existing, proposed, and future characters across functions and time periods.

1 Summary

This document proposes encoding some numerals used in the Uruk and Early Dynastic periods in conjunction with the Sumero-Akkadian cuneiform script¹ and the proto-cuneiform script². The proposed characters are listed in section 2.

The non-numeric signs of proto-cuneiform will be the subject of a separate proposal; we need only note here that the divergence between the approaches to character identity in modern scholarship requires that proto-cuneiform be disunified from cuneiform: proto-cuneiform is effectively treated as an undeciphered script. In contrast, the cuneiform encoding model is semantic, requiring an understanding of the text to correctly encode it.

However, the *numerals* used in proto-cuneiform should be unified with ones used in the Early Dynastic period, for the reasons set forth in section 4. The proposed “curved”, or “curviform”, numerals³ should however *not* be unified with the already-encoded cuneiform numerals⁴. Since the encoding proposals for the cuneiform script twenty years ago provisionally considered the curviform numerals to be glyph variants of the cuneiform numerals, a detailed rationale is provided in section 3, including compatibility considerations in section 3.7.

The overall picture of unifications and disunifications over time is illustrated in table 1. The Script_Extensions property assignments in section 2.2 reflect the overlap.

[TODO(egg): Mention the other sections here too.]

2 Proposed changes to the Standard

2.1 Summary of proposed characters

2.2 Properties

2.3 Character names list

2.4 Core specification text

3 Rationale for curviform–cuneiform disunification

TODO(egg): blurb.





¹ISO 15924: Xsux, Script property value long name: Cuneiform; encoded since Unicode Version 5.0.

²ISO 15924: Pcun, not yet encoded.

³𐎶 1-9(aš^c = N_1), 𐎵 1-5(u^c = N_{14}), 𐎶𐎵 1-9(𐎶š₂^c = N_{34}), 𐎶𐎵𐎶 1-5(𐎶š^cu^c = N_{48}), etc.

⁴𐎶𐎵𐎶 1-9(aš), 𐎶𐎵𐎶 1-5(u), 𐎶𐎵𐎶 1-9(𐎶š₂), 𐎶𐎵𐎶 1-5(𐎶š^cu), etc.

3.1 The cuneiform encoding model



















As outlined in, *e.g.*, [UTR56], the cuneiform encoding model is diachronic; each character may have wildly different glyphs depending on time period and region. For instance, the sign IM may resemble  in texts from Early Dynastic IIIa Šuruppag as in the character code charts,  later in the third millennium⁵,  in Old Babylonian cursive,  in Neo-Assyrian, but is always encoded as U+1214E CUNEIFORM SIGN IM.


This encoding model allows for the interoperable representation of editions of diachronic reference works such as sign lists⁶ and dictionaries⁷, and of composite texts⁸. By being compatible with similarly diachronic transliteration practice (that is, by avoiding distinctions finer than those made in transliteration), the encoding model also allows for automated conversion of transliterated corpora to cuneiform, which has proven useful as a processing step in analyses such as [Rom24; JJ24]⁹. The diachronic approach is also useful for pedagogical applications¹⁰.

3.2 Arguments for curviform–cuneiform unification

In this context, the argument was made in [Ando4], as part of discussion of the cuneiform encoding¹¹ that the curviform numerals, which occasionally appear in the Ur III period and are used heavily in the Early Dynastic period, were a stylistic distinction unifiable with the cuneiform digits, and that an archaizing Ur III font or an Early Dynastic font could have curviform glyphs for the appropriate characters.







































Some co-occurrence of curviform and cuneiform digits was known and acknowledged. [Ando4, p. 3] cites [NDE93, p. 62], which is a copy of [P020054], an Early Dynastic IIIb administrative tablet from Nirsu. The excerpt cited, lines 1–3 of column 1 of the obverse, is as follows:

						
1(NEŠ ₂)	1(U)	1/2(DIŠ)	5(DIŠ <i>tenû</i>)	gi	us ₂	sa ₂
	7.5 (ropes)		5	reed	side	equal
						
3(U)	6(DIŠ <i>tenû</i>)	gi	saṇ	sa ₂		
3(ropes)	6	reed	front	equal		
						
ašag-bi	1(BUR ₃ ^c)	1(EŠE ₃ ^c)	1(IKU ^c)	1/2(IKU ^c)		
this field						

⁵Merging with U+1224E  NI₂.

⁶Notably [VT+14] and the online edition of [Bor10] in [Jim+23, Signs].

⁷Notably [TJV17] and the online edition of [Sch10] in [Jim+23, Dictionary].

⁸For example, there are Neo-Assyrian and Neo-Babylonian copies parts of the laws of                                      

𐎶𐎵𐎶𐎵𐎶𐎵𐎶𐎵
 tug_x(LAK483)-si-ga-kam¹⁴
 deep ploughing

The argument made in [Ando4, p. 4] is that this is comparable to a stylistic distinction such as¹⁵

465 metres, equal lengths
 198 metres, equal widths
 this field: 9, 18 hectares, deeply ploughed

where the numerals have the same structure ([Ando4] contrasts this to the different structures of ASCII digits and roman numerals). That document further claims that “the number signs do not normally carry in their individual signs the meaning of what they are used to measure”, and that curviform and cuneiform numerals “are not normally mixed together in a single numerical expression”. In addition, [Ando4, p. 4] points out that the cuneiform numeric signs are descended from the curviform ones (this is undisputed), and claims there is only a small re-allocation of the function of signs (from 𐎶- to 𐎵- numerals). It therefore comes to the conclusion that the use of curviform numerals should be seen as a formatting distinction, rather than one that should be represented in plain text, and insists that the encoding should capture the lineal historical descent of those signs, presumably to take advantage of the benefits of diachronic encoding described in section 3.1.

Although they had been part of the preliminary proposal [EFT03], the curviform numerals were therefore removed from [EFT04b] and [EFT04a], which both state that “The distinction between curved numerals and their cuneiform descendants is treated as glyphic for the purposes of the present proposal; this issue will need to be revisited in subsequent encoding phases.”

The time has come to revisit this issue. As we will see in section 3.3, numerals can only be interpreted in the context of what they measure *i.e.*, as part of a metrological system. In section 3.4 we will see that in some periods:

- the functions and use of the numerals vary beyond the mere 𐎶/𐎵 switch;
- the contrast between curviform and cuneiform numerals is commonly used to distinguish metrological systems;
- some metrological systems commonly mix curviform and cuneiform in single numerical expressions.

3.3 Metrology

𐎶𐎵 𐎶𐎵𐎶𐎵𐎶𐎵 𐎶𐎵
 𐎶𐎵 𐎶𐎵 𐎶𐎵𐎶𐎵 𐎶𐎵 𐎶𐎵
 𐎶𐎵 𐎶𐎵𐎶𐎵 𐎶𐎵 𐎶𐎵 𐎶𐎵𐎶𐎵

I want to write tablets: the tablet of 1 gur of barley to 600 gur; the tablet of 1 shekel of silver to 10 minas [...]

Edubba'a D

Before diving into the usage of the curviform numerals in the Early Dynastic period to explain the contrast with cuneiform numerals, it is useful to understand the usage of the already-encoded characters in the Ur III and Old Babylonian periods.






¹⁴Transliteration after [Lec20, p. 8].

¹⁵We have taken the liberty of adjusting the analogy to use measures approximately equal to those in [P020054], instead of a field of five by twenty-five metres.








- <- used in $G_{Ur III/OB}$, of which <- are also used in $S_{Ur III/OB}$ and the SPVS as well as with overt units;
- | used in $S_{Ur III/OB}$, and sometimes with overt units;
- | used in $S_{Ur III/OB}$;
- ◇ used in $S_{Ur III/OB}$ and $G_{Ur III/OB}$;
- ◇ used in $S_{Ur III/OB}$ and $G_{Ur III/OB}$;
- — used in $C_{Ur III/OB}$ as well as with overt units of the weight system;
- +, +, +, +, + used in $C_{Ur III/OB}$;
- |, |, |, | used in $C_{Ur III/OB}$ —note the overlap with | used in $S_{Ur III/OB}$;
- < and < used in $G_{Ur III/OB}$.

3.3.4 The length system

In the Ur III and Old Babylonian periods, lengths are expressed using overt units counted with | and <- numerals with their system $S_{Ur III/OB}$ values²⁴. Since it does not have any unusual numerals, this system would not in itself be of much relevance to character encoding, but we present it here as background for its Early Dynastic counterpart presented in section 3.4. Metrological tables use the following units [Fri07, p. 118; Rob19]:

	← 60		← 10		← 12		← 30		($L_{Ur III/OB}$)
danna		US ²⁵		nindan		kuš ₃		šu-si	
bērum				nindanum		ammatum		ubānum	
league		cable		rod		cubit		finger	
10,8 km		360 m		6 m		50 cm		17 mm	

Two more units appear occasionally [Pow87, p. 459; Fri07, p. 118; Rob19]:

	← 30		← 6		← 10		← 2		← 6		← 30		($\bar{L}_{Ur III/OB}$)
eše ₂		ašlum		qānum		gi		reed		finger			
		rope						3 m					
		60 m											

In addition, there are Akkadian names for the half-rope and half-reed, see [Pow87, pp. 463 sq.].

3.3.5 Fractions

TODO

3.4 Early metrology

At first sight, the metrological systems from the Early Dynastic period match the ones previously mentioned. In particular, the discrete counting system used in the Early Dynastic period (and earlier in the Uruk period) clearly mirrors system $S_{Ur III/OB}$ [Fri07, p. 374; DE87, pp. 127, 165]:

$$\odot \xleftarrow{10} \bullet \xleftarrow{6} \text{D} \xleftarrow{10} \text{D} \xleftarrow{6} \bullet \xleftarrow{10} \text{D}. \quad (S)$$

²⁴Adjacent units are no more than a factor of 60 apart, so higher numerals such as | or ◇ are not used.

Likewise the area system used in the Early Dynastic IIIb period mirrors system $G_{Ur III/OB}$ [Deimel1922; NDE93, p. 63; Fri07, p. 378; Gom16]:

$$\odot \xleftarrow{10} \bullet \xleftarrow{6} \star \xleftarrow{10} \bullet \xleftarrow{3} \text{𐎶} \xleftarrow{6} \text{𐎶}, \quad (G_{ED IIIb})$$

As noted in [Ando4, p. 4] (see section 3.2), the vertical \uparrow from $S_{Ur III/OB}$ becomes a horizontal 𐎶 in system S . It is however far from the only case of such a reallocation of function. The earlier form of System G was [DE87, pp. 141, 165; Fri07, p. 378]:

$$\bullet \xleftarrow{6} \odot \xleftarrow{10} \bullet \xleftarrow{3} \text{𐎶} \xleftarrow{6} \text{𐎶}, \quad (G)$$

Observe that, as noted in [DE87, p. 142], \odot changes meaning from $10\bullet$ in system G to $10\bullet$ in system $G_{ED IIIb}$. System G is used in the Uruk period, but also in the ED I–II period (it is the “area 2” system in [Chambon2003], whereas $G_{ED IIIb}$ is the “area 1” system).

The length system Early Dynastic IIIb of the state of Lagaš is of particular interest. As described in [Pow87, p. 466; Lec20, pp. 289 sq.], lengths are expressed in rods, but the unit sign 𐎶 is generally omitted; in addition, only ten of rods are used; these are equal to one rope, but the sign 𐎶 is not written either. Length shorter than one rope are expressed in half-rope using the $1/2$ sign 𐎶 (again with no 𐎶), and then in reeds, *with* the sign 𐎶 . Effectively, this yields the following factor diagram:

$$\uparrow \xleftarrow{6} \text{𐎶} \xleftarrow{2} \text{𐎶} \xleftarrow{10} \text{𐎶} \text{𐎶}. \quad (L_{ED IIIb})$$

eše_2
 $ašlum$
 rope
 60 m

gi
 $qānum$
 reed
 3 m

This is the system that was used to express the sides of the field in [P020054] discussed in section 3.2. In that tablet and others from the same period, such as the ones, areas are expressed in system $G_{ED IIIb}$, with curviform numerals²⁶; in the absence of overt units, the use of curviform or cuneiform numerals therefore disambiguates a numeric expression between an area and a length, and therefore the interpretation of its numerals between systems $G_{ED IIIb}$ and $L_{ED IIIb}$. The sign 𐎶 , which would also disambiguate the interpretation as an area, is sometimes used after areas in ED IIIb Lagaš, but not systematically; in particular the area of the first field in [P020054] does not use this suffix. See [Lec20] for many examples with and without 𐎶 .

[Pow87, p. 466] notes that reeds “are regularly written with the normal, cuneiform end of the stylus. Higher units are usually written with the reversed (round) end of the stylus.” Powell does not elaborate on the specifics of this mixed use of numerals, but a cursory search in CDLI finds many occurrences: [cite VS 14, 100, 130, 74, VS25, 83... Just search the downloaded ATF pile for @c.*disz@t\) gi] From these we can draw the following factor diagram:

²⁶TODO(egg): Note the handful of late Urukagina tablets that start to have cuneiform areas.

3.4.1 Use in modern publications

3.5 Non-numeric usage

[illegible]

The beginning of the scribal art is a single wedge. That one has six pronunciations; it also stands for 'sixty'. Do you know its reading?

Examenstext A

3.6 Limited benefits of diachronic encoding for numerals

[Composite texts dating back to the period where curved numerals are in use tend to be limited to lexical texts, which do not usually have numbers. When they do, diachronic encoding is prevented by diš-aš distincticons anyway. Administrative texts, which are where numbers are most prominent, are not composite.]

[Diachronic reference works tend to not include numbers, or when they do, to treat them specially (for instance, they are shown at the end of sign lists such as `TODO`).]

[The overarching goal of having consistent representation for equivalent numeric expressions from different periods is quickly foiled by changes in metrology.]

Note that in [Rom24] [TODO(egg): Cite the GitHub repository], as in many other such analyses, numbers are removed as an early step in processing; these therefore would not benefit from diachrony in the encoding of numeric expressions.

3.6.1 Compatibility with transliteration

3.7 Compatibility considerations

3.7.1 The case of ŠAR₂

4 Rationale for ED-Uruk numeral unification

5 Considerations on individual numeral series

[TODO Document to the extent possible the metrological systems in which each sign is used. Note the disunification of N9 and N10 from 4(ban₂@c) and 5(ban₂@c).]

6 Characters not included in this proposal

6.1 Missing numerals

(N_{17} , $12N_{14}$, etc.) 7(diš *tenû*)

6.2 Stacking patterns

(... are a mess, vary within Uruk, and are not transliterated/documented by Englund, so let's not go there for now.)

6.3 Matters for higher-level protocols

Rotated bits: <https://cdli.mpiwg-berlin.mpg.de/artifacts/101087>

7 Acknowledgements

References

- [Ando4] L. Anderson. *Unification of Cuneiform Numbers*. 2004.
UTC: [L2/04-099](#).
- [Bor10] R. Borger. *Mesopotamisches Zeichenlexikon*. Alter Orient und Altes Testament 305. Ugarit-Verlag, 2010.
- [Cap02] R. Caplice. *Introduction to Akkadian*. 4th ed. Editrice Pontificio Istituto Biblico, 2002.
ISBN: 88-7653-566-7.
- [Con24] P. Constable, ed. *Minutes of UTC Meeting 180* (July 23–25, 2024). July 29, 2024.
UTC: [L2/24-159](#).
- [DE87] P. Damerow and R. K. Englund. “Die Zahlzeichensysteme der archaischen Texte aus Uruk”. In: M. W. Green and H. J. Nissen. *Zeichenliste der archaischen Texte aus Uruk*. Archaische Texte aus Uruk 2. An offprint of this chapter is available at <https://cdli.mpiwg-berlin.mpg.de/files-up/publications/englund1987a.pdf>. Gebr. Mann Verlag, 1987. Chap. 3, pp. 117–165.
- [EF03] M. Everson and K. Feuerherm. *Basic principles for the encoding of Sumero-Akkadian Cuneiform*. May 25, 2003.
UTC: [L2/03-162](#).
- [EFT03] M. Everson, K. Feuerherm, and S. Tinney. *Preliminary proposal to encode the Cuneiform script in the SMP of the UCS*. Nov. 3, 2003.
UTC: [L2/03-393R](#).
- [EFT04a] M. Everson, K. Feuerherm, and S. Tinney. *Final proposal to encode the Cuneiform script in the SMP of the UCS*. June 8, 2004.
UTC: [L2/04-189](#).
- [EFT04b] M. Everson, K. Feuerherm, and S. Tinney. *Revised proposal to encode the Cuneiform script in the SMP of the UCS*. Jan. 29, 2004.
UTC: [L2/04-036](#).
- [För16] W. Förtsch. *Altbabylonische Wirtschaftstexte aus der Zeit Lugalanda’s und Urukagina’s*. Vorderasiatische Schriftdenkmäler der Königlichen Museen zu Berlin 14. J. C. Hinrichs, 1916.
- [Fri07] J. Friberg. *A Remarkable Collection of Babylonian Mathematical Texts. Manuscripts in the Schøyen Collection: Cuneiform Texts I*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, 2007.
ISBN: 978-0-387-34543-7.

- [Fri78] J. Friberg. *A Method for the Decipherment, through Mathematical and Metrological Analysis, of Proto-Sumerian and Proto-Elamite Semi-Pictographic Inscriptions*. The Third Millenium Roots of Babylonian Mathematics 1. Department of Mathematics, Chalmers University of Technology, 1978.
- [Gom16] B. Gombert. “ED IIIb metrology: texts from Lagaš”. In: *CDLI:wiki. A Library of Knowledge of the Cuneiform Digital Library Initiative*. Apr. 12, 2016.
eprint: https://cdli.ox.ac.uk/wiki/doku.php?id=ed_iii_metrological_systems.
- [Hue11] J. Huehnergard. *A Grammar of Akkadian*. 3rd ed. Brill, 2011.
ISBN: 978-1-57506-941-8.
- [Jim+23] E. Jiménez, Z. Földi, A. Härtinen, A. Heinrich, T. Mitto, G. Rozzi, I. Khait, J. Laasonen, F. Simonjetz, et al., eds. *electronic Babylonian Library*. 2023–.
eprint: <https://www.ebl.lmu.de/>.
- [JJ24] T. Jauhainen and H. Jauhainen. “Advancing Cuneiform Text Dating Through Automatic Analysis”. 69th Rencontre Assyriologique Internationale (July 8–12, 2024). July 11, 2024 14:00.
- [Lec20] C. Lecompte. “The Measurement of Fields During the Pre-sargonic Period”. In: *Mathematics, Administrative and Economic Activities in Ancient Worlds*. Ed. by C. Michel and K. Chemla. Why the Sciences of the Ancient World Matter 5. Springer, 2020.
- [NDE93] H. J. Nissen, P. Damerow, and R. K. Englund. *Archaic Bookkeeping. Early Writing and Techniques of Economic Administration in the Ancient Near East*. Trans. by P. Larsen. The University of Chicago Press, 1993.
ISBN: 0-226-58659-6.
- [Oel22] J. Oelsner. *Der Kodex Hammu-rāpi*. dubsar 4. Zaphon, 2022.
- [P020054] VAT 04731. [För16, 40 p.14]. Vorderasiatisches Museum.
CDLI: [P020054](#).
- [P102305] X.3.139. Michael C. Carlos Museum, Emory University.
CDLI: [P102305](#).
- [Pow87] M. Powell. “Maße und Gewichte”. In: *Reallexikon der Assyriologie und vorderasiatischen Archäologie*. Ed. by D. O. Edzard. Vol. 7 Libanukšabaš-Medizin. 1987–1990, pp. 457–530.
- [Rob08] E. Robson. *Mathematics in Ancient Iraq. A Social History*. Princeton University Press, 2008.
ISBN: 978-0-691-09182-2.
- [Rob19] E. Robson. “Oracc metrology guidelines”. In: *Oracc: The Open Richly Annotated Cuneiform Corpus*. Dec. 18, 2019.
eprint: <http://oracc.org/doc/help/editinginatf/metrology/metrologicaltables/>.
- [Rob22] E. Robson. “Overview of Metrological Systems”. In: *The Digital Corpus of Cuneiform Mathematical Texts*. 2022.
eprint: <http://oracc.org/dccmt/Metrology/>.

- [Rom24] A. Romach. “The Neo Assyrian Land Sale Documents from Dur-Katlimmu: A Stylometric Analysis of Their Scribal Features”. 69th Rencontre Assyriologique Internationale (July 8–12, 2024). July 10, 2024 12:00.
- [Sch10] W. Schramm. *Akkadische Logogramme*. Göttinger Beiträge zum Alten Orient 5. Universitätsverlag Göttingen, 2010.
ISBN: 9783941875654.
DOI: [10.17875/gup2010-511](https://doi.org/10.17875/gup2010-511).
- [Svä+24] S. Svärd, M. Lorenzon, J. Töyräänvuori, J. Valk, T. Alstola, E. Bennett, R. Uotila, and T. Auranne, eds. *RAI 69 Abstracts*. July 2024.
eprint: https://www.helsinki.fi/assets/drupal/2024-07/RAIAbstractBookAjoitettuJaPäiväty_1.pdf.
- [TJV17] S. Tinney, P. Jones, and N. Veldhuis, eds. *The electronic Pennsylvania Sumerian Dictionary*. 2nd ed. 2017–.
eprint: <http://oracc.org/epsd2>.
- [Uni16] The Unicode Consortium. *The Unicode Standard*. Version 16.0.0. The Unicode Consortium, Sept. 10, 2024.
ISBN: 978-1-936213-34-4.
eprint: <https://www.unicode.org/versions/Unicode16.0.0/core-spec/>.
- [UTR56] R. Leroy, ed. *Unicode Cuneiform Sign Lists*. Unicode Technical Report #56.
eprint: <https://www.unicode.org/reports/tr56/>.
- [VT+14] N. Veldhuis, S. Tinney, et al., eds. *Oracc Sign List*. 2014–.
eprint: <http://oracc.org/osl/>.