

Archaic cuneiform numbers

Robin Leroy, Anshuman Pandey, and Steve Tinney

2024-08-24

Contents

1	Summary	2
2	Proposed changes to the Standard	3
2.1	Summary of proposed characters	3
2.2	Properties	3
2.3	Character names list	3
2.4	Core specification text	3
3	Rationale for curviform–cuneiform disunification	3
3.1	The cuneiform encoding model	3
3.2	Arguments for curviform–cuneiform unification	3
3.3	A primer on classic Ur III and Old Babylonian metrologies	5
3.3.1	The discrete counting system	5
3.3.2	The area system	6
3.3.3	The capacity system	6
3.3.4	The length system	7
3.3.5	Fractions	8
3.4	Curviform numerals in early metrologies	8
3.4.1	Field lengths in Nirsu	8
3.4.2	Dyke lengths in Nirsu	9
3.4.3	Grain in Nirsu	9
3.4.4	Grain in Ebla	9
3.4.5	Use in modern publications	11
3.5	Non-numeric usage	13
3.6	Limited benefits of diachronic encoding for numerals	14
3.7	Compatibility considerations	15
3.7.1	The case of ŠAR ₂	15
3.7.2	Transliteration	16
3.8	Conclusions	17
4	Rationale for ED–Uruk numeral unification	18
5	Considerations on individual numeral series	18
6	Characters not included in this proposal	18
6.1	Missing numerals	18
6.2	Stacking patterns	18

Acknowledgements	19
References	20
Artefacts	20
Unicode documents	21
Other documents	22

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.]

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

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

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

²ISO 15924: Pcun, not yet encoded.

³Impressed into clay using cylindrical styli, held either perpendicular to the tablet, yielding • (small stylus) or ● (large stylus), or at a shallower angle: ◻, ◻◻ (small stylus), ◻◻◻ (large stylus). Some numerals are composed of multiple such impressions, e.g., ◻◻◻◻.

⁴Impressed into clay using a stylus with a trihedral end: ◻ (stylus held horizontally), ◻◻ (vertically), ◻◻◻ (diagonally) ◻◻◻◻ (diagonally with the stylus rotated along its axis), ◻◻◻◻◻ (stylus pressed deeper, forming a larger wedge), ◻◻◻◻◻◻ (combining ◻◻◻◻ and ◻◻◻◻), etc.

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.

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 millenium⁵,  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, *i.e.*, 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 pedagogic applications¹⁰.

3.2 Arguments for curviform–cuneiform unification















































































































































































































































































































































































In this context, the argument was made in [L2/04-099], 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. [L2/04-099, 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:

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                                                         

                                                          

                                                          

                                                          

                                                          

                                                          

             

 ¹²						
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
 ¹³						
3(U)	6(DIŠ <i>tenû</i>)	gi	saṇ	sa ₂		
3(ropes)	6	reed	front	equal		
						
ašag-bi	1(BUR ₃)	1(EŠE ₃)	1(IKU ^c)	1/2(IKU ^c)		
this field						
						
				tug _x (LAK483)-si-ga-kam ¹⁴		
				deep ploughing		



The argument made in [L2/04-099, 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 ([L2/04-099] 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”, noting the exceptions of [P232278; P232280]. In addition, [L2/04-099, 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  to  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 [L2/03-393R], the curviform numerals were therefore removed from [L2/04-036] and [L2/04-189], 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 / switch;

¹²As noted in [Pow87, p. 466], this sign has a very short “tail” in this period, so that it is wider than it is tall, and can at first seem like a large  in copies. The photos in CDLI clearly show that this is in fact a vertical wedge.

¹³Note that ED IIIb  numerals have a somewhat different appearance from those of the Ur III period used in this transcription; the sign  in [P020054] looks more like Ur III .

¹⁴Transliteration after [Lec20, p. 325].

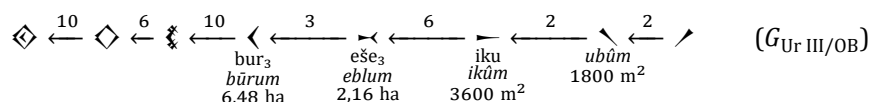
¹⁵We have taken the liberty of adjusting the analogy to use measures approximately equal to those in [P020054], instead of a field of five by twenty-five metres.










- ### 3.3 A primer on classic Ur III and Old Babylonian metrologies



3.3.2 The area system

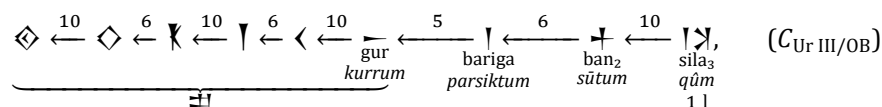
The discrete counting system was not the only non-positional system in use in the Ur III and Old Babylonian periods; different systems were in use depending on what was being counted or measured. For instance, field areas were measured using the following system, where for the named units we have provided the name of the unit in transliterated Sumerian, normalized Old Babylonian Akkadian, and the approximate metric equivalent [Fri07, p. 378; Rob19]:








Note that for the range of areas given above¹⁸, this system does not use any symbols separate from the numerals for the individual units (*ubûm*, *ikûm*, *ebum*, and *bûrum*). As mentioned in [Rob19], the whole numeric expression for the area would be followed by the sign  functioning as punctuation¹⁹, but the numerals are tied to the metrology; thus a surface of 5 *bûr* 1 *ebel* 4 *ikû* (100 *ikû*, 36 ha) would be written²⁰     . Contrast this with systems where the same numerals are used for different units, and overt units are used, as in “88 acres 3 roods 33 perches” or 五頃八畝五分九厘. Note also that the same signs are shared between multiple systems, with different relations; the sign  is equal to sixty times  in the area system, but to three hundred and sixty times  in the discrete counting system.

3.3.3 The capacity system

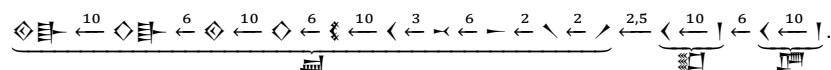
Another such system of note is the one for capacities²¹ [Fri07, p. 376; Rob19],



where the numerals for ban₂ are 𐤁, 𐤂, 𐤃, 𐤄, and 𐤅, and those for bariga are 𐤆, 𐤇, 𐤈, and 𐤉 (contrast ordinary 𐤈 and 𐤉 otherwise used with 𐤆 numerals). As described in [Hue1, p. 585 with notes (b) and (f)], the sign 𐤆𐤆 GUR, while it is used only with volumes in excess of one gur, is written after the whole expression, after the overt unit sign 𐤇 if present, and after the word for “grain” if present, as in

 3554 gur
  3 ban₂
  6
  sila₃
  of grain.

¹⁸For areas smaller than a quarter *ikûm*, an overt unit is used, with 1 *mûšarum* (36 m²) written 𒄠, equal to one hundredth of an *ikûm*, then sexagesimally subdivided in 60 𒄡 (shekels). For areas greater than 3600 *būrû*, the 𒄠 and 𒄡 numerals are reused with a suffix 𒄢 (gal, Sumerian: big), as follows [Robo8, p. 295 with notes b and c; Fri07, p. 378; Rob19]:



¹⁹TODO(egg): acknowledge Proust 2020 but note that this is irrelevant to encoding concerns






²⁰As in the surface of the field of (the city of Apisal) reported on [P102305, r. 1]


²¹Used for volumes of grain, but also oil, dairy products, beer, etc., as well as to express the capacity of boats; volumes of earthworks instead use system $G_{U_{rIII}/OB}$ based on a height of one cubit, see [Pow87, p. 488; Rob08, p. 294; Rob19].

This intertwining of units and numerals explains the large number of already-encoded numeral series:

- \mathbb{I} - \mathbb{W} used in $S_{\text{Ur III/OB}}$ and the SPVS as well as with overt units;
- \mathbb{G} - \mathbb{X} used in $G_{\text{Ur III/OB}}$, of which \mathbb{G} - \mathbb{X} are also used in $S_{\text{Ur III/OB}}$ and the SPVS as well as with overt units;
- \mathbb{I} - \mathbb{W} used in $S_{\text{Ur III/OB}}$, and sometimes with overt units;
- \mathbb{K} - \mathbb{X} used in $S_{\text{Ur III/OB}}$;
- \mathbb{D} - \mathbb{X} used in $S_{\text{Ur III/OB}}$ and $G_{\text{Ur III/OB}}$;
- \mathbb{E} - \mathbb{X} used in $S_{\text{Ur III/OB}}$ and $G_{\text{Ur III/OB}}$;
- \mathbb{T} - \mathbb{X} used in $C_{\text{Ur III/OB}}$ as well as with overt units of the weight system;
- \mathbb{T} , \mathbb{T} , \mathbb{T} , \mathbb{T} , \mathbb{T} used in $C_{\text{Ur III/OB}}$;
- \mathbb{I} , \mathbb{I} , \mathbb{W} , \mathbb{W} used in $C_{\text{Ur III/OB}}$ —note the overlap with \mathbb{I} - \mathbb{W} ;
- \mathbb{X} and \mathbb{X} used in $G_{\text{Ur III/OB}}$.

3.3.4 The length system

 60 ←  10 ←  12 ←  30 ←  (L_{Ur III/Ob})
 danna
berum
league
10.8 km
 US²⁵
cable
360 m
 nindan
nindanum
rod
6 m
 kuš₃
ammatum
cubit
50 cm
 šu-si
ubānum
finger
17 mm


($\bar{l}_{Ur III/OB}$)

eš₂
ašlum
 rope
 60 m

gi
qānum
 reed
 3 m

²⁵TODO

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 Curviform numerals in early metrologies

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]:

$$\odot \xleftarrow{10} \bullet \xleftarrow{6} \text{◐} \xleftarrow{10} \text{◑} \xleftarrow{6} \bullet \xleftarrow{10} \text{◒}. \quad (S)$$

Likewise the area system used in the Early Dynastic IIIb period mirrors system $G_{Ur III/OB}$ [Dei22, p. 72; NDE93, p. 63; Fri07, p. 378; Gom16]:

$$\odot \xleftarrow{10} \bullet \xleftarrow{6} \text{⊛} \xleftarrow{10} \bullet \xleftarrow{3} \text{◐} \xleftarrow{6} \text{◒}, \quad (G_{ED IIIb})$$

As noted in [L2/04-099, p. 4] (see section 3.2), the vertical \uparrow from $S_{Ur III/OB}$ becomes a horizontal ◒ in system S . It is however far from the only case of such a reallocation of function. The earlier form of System G is [DE87, pp. 141, 165; Fri07, p. 378]:

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

Observe that, as noted in [DE87, p. 142], \odot changes meaning from $10\bullet$ in system G to $600\bullet$ 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 [Cha03], whereas $G_{ED IIIb}$ is the “area 1” system).

3.4.1 Field lengths in Nirsu

The length system of the Early Dynastic IIIb 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 𒌦 is generally omitted; in addition, only tens of rods are used; these are equal to one rope, but the sign 𒌦 is not written either. Length shorter than one rope are expressed in half-rope using the $1/2$ sign 𒌦 (again with no 𒌦), and then in reeds, *with* the sign 𒌦 . Effectively, this yields the following factor diagram:

$$\begin{array}{c} \uparrow \xleftarrow{6} \text{◐} \xleftarrow{2} \text{𒌦} \xleftarrow{10} \text{𒌦} \text{ }^{28} \\ \begin{array}{l} 1 \text{ eše}_2 = 10 \text{ nindan} \\ 1 \text{ rope} = 10 \text{ rods} \\ 60 \text{ m} \end{array} \qquad \begin{array}{l} \text{gi} \\ \text{reed} \\ 3 \text{ m} \end{array} \end{array} \quad (L_{ED IIIb})$$

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 discussed in [Lec20], areas are expressed in system $G_{ED IIIb}$, with curviform numerals²⁹; in the absence of overt units, such as when dealing with length that

²⁸Note that the reeds are counted using *tenû* numerals, 𒌦 , 𒌦 , 𒌦 , etc.

²⁹TODO(egg): Note the handful of late Urukagina tablets that start to have cuneiform areas.

3.4.2 Dyke lengths in Nirsu

— [P221305, obv. 1, 4]^[32] 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎

— [P020129, rev. 2, 1] 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎

— [P221291, rev. 5, 1]^[33] 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎

— [P221266, rev. 2, 1] 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎 𠄎

$$\underbrace{\begin{array}{c} 10 \\ \leftarrow \\ \text{Diagram 1} \end{array}}_{\text{Diagram 2}} \leftarrow 6 \bullet = \underbrace{\begin{array}{c} 2 \\ \leftarrow \\ \text{Diagram 3} \end{array}}_{35} \leftarrow 10 \leftarrow 6 \leftarrow 3 \leftarrow \text{Diagram 4} \quad (L'_{\text{ED IIIb}})$$

3.4.3 Grain in Nirsu

3.4.4 Grain in Ebla

³⁰This is the case of the sides of the field in [P020054, obv. ii 2–3].

³²CDLI only has a copy, but a photo may be found in [Lec12, p. 82]. On that photo the 𒀭𒀭𒀭𒀭 is not visible. Lecompte notes that the copy is faithful; indeed another 𒀭𒀭𒀭𒀭 can be seen both on the copy and the photo on obv. 2, 2.

³³From copy.

³⁴TODO Cite also DP 568, the one with and even though it has no reeds.

³⁵TODO(egg): Note that one unit may be omitted if the other is present

The system of grain³⁶ capacities in Ebla uses the following units³⁷:

$$\begin{array}{c} \text{gu}_2\text{-bar} \xleftarrow{2} \text{ba-ri}_2\text{-zu} \xleftarrow{\frac{5}{2}} \text{nin}_4 \xleftarrow{4} \text{nin}_2\text{-sagšu} \xleftarrow{6} \text{an-zam}_x \end{array}$$

The $\text{gu}_2\text{-bar}$ and $\text{ba-ri}_2\text{-zu}$ are generally counted using curviform numerals, and the smaller units using cuneiform nin numerals. Indeed, a search on [Mil+07] for co-occurrences of an-zam_x with either of $\text{gu}_2\text{-bar}$ or $\text{ba-ri}_2\text{-zu}$ finds the following expressions³⁸:

1. [P240532, verso 4, 9] $\text{gu}_2\text{-bar} \text{ ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$ ⁴⁰
2. [P240548, verso 1, 1] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$
3. [P240655, recto 7, 9] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$ ⁴¹
4. [P240579, verso 4, 3] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$
5. [P240675, verso 2, 2] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$
6. [P240609, verso 3, 1] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$
7. [P240533, recto 3, 3] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x \text{ an-zam}_x$
8. [P240697, recto 1, 5] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$ ⁴²
9. [P240653, recto 6, 2] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x \text{ an-zam}_x$
10. [P240654, recto 2, 6] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x \text{ an-zam}_x$ ⁴³
11. [P240531, recto 1, 8] $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x \text{ an-zam}_x$
12. [P241708, recto 1, 1]⁴⁵ $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$
13. [P241904, recto 1, 1]⁴⁶ $\text{ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$ ⁴⁷

Note that higher numbers of $\text{gu}_2\text{-bar}$ are expressed in hundreds (*mi-at* mi-at) and then thousands (*li-im* li-im), as is typical in Ebla [Arc15, p. 33], e.g., in [P240532, verso 2, 3], $\text{mi-at} \text{ li-im} \text{ ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x$ (100 + 60 + 30 + 5 = 195 $\text{gu}_2\text{-bar}$ of grain).

³⁶Liquid capacities use a different system [Arc15, p. 229 with note 12]:

$$\text{la-ha} \xleftarrow{30} \text{sil}_3 \xleftarrow{6} \text{an-zam}_x$$

At a glance it seems that sil_3 are counted with cuneiform numerals and higher units with curviform ones, thus

$$\text{mi-at} \text{ li-im} \text{ ba-ri}_2\text{-zu} \text{ nin}_4 \text{ an-zam}_x \xleftarrow{\frac{5}{2}} \text{ba-ri}_2\text{-zu} \xleftarrow{6} \text{ba-ri}_2\text{-zu} \xleftarrow{10} \text{ba-ri}_2\text{-zu} \xleftarrow{3} \text{ba-ri}_2\text{-zu} \xleftarrow{\frac{10}{3}} \text{ba-ri}_2\text{-zu} \xleftarrow{6} \text{ba-ri}_2\text{-zu}$$

but we have not investigated this thoroughly.

³⁷TODO mention the other one citing Chambon and the footnote in Archi

³⁸We cite here only one attestation per tablet; most tablets contain several expressions mixing curviform $\text{ba-ri}_2\text{-zu}$ and larger with cuneiform nin and smaller. In all cases the transcriptions given here are based on the EbDA transliterations, but the shape and orientation of the numerals was checked³⁹ on a photograph (from EbDA unless noted otherwise).

³⁹As we will see in Section 3.7.2, CDLI transliterations indicate numeral shape; however, as of this writing, they do so incorrectly on the Ebla corpus, claiming that all numerals are curviform, so we were not able to rely on them in this specific case.

⁴⁰ba-ri₂-zu₂, a variant spelling.

⁴¹Short for $\text{gu}_2\text{-bar}$.

⁴²Note the omitted $\text{gu}_2\text{-bar}$.

⁴³Instead of the expected $\text{ba-ri}_2\text{-zu}$.

⁴⁴ an-zam_x not legible on the EbDA photo.

⁴⁵From CDLI photo.

⁴⁶From photo in [Arc89, p. 6].

⁴⁷Laid out as $\text{ba-ri}_2\text{-zu}$; on stacking patterns see Section 6.2.

TODO figure

Figure 4: TODO [Cha03, p. 6]

formed by only two signs Υ and ◀ , repeated as many times as necessary; this type of notation is highly standardized. Second, the order of magnitude of the numbers noted in this system is not indicated: 1, 60, 60², 60³, 1/60, 1/60², etc. are written in the same way, with the vertical wedge Υ . The third feature concerns the exact function of

Figure 5: TODO [Cha12, p. 58]

one step. The scribes of the Early Dynastic Period (c. 2600 BC), for instance, represented the number 648,000 with: $\Upsilon\text{◀}\bullet\bullet\bullet$ but never with the repetition $\Upsilon\text{◀}\Upsilon\text{◀}\Upsilon\text{◀}\bullet\bullet\bullet$.

Figure 6: TODO [Cha12, p. 59]⁵⁰

repetition of the same sign refers to both the capacity unit signified—often but not necessarily written immediately afterwards—and its value. The units of measurement are written in descending order from left to right—just as we would write 3 km, 120 m, 50 cm. For example:

$\Upsilon\text{◀}\text{◀}\text{◀}$ še bar ◀ ba-rí-zu
 ‘3 gubar (capacity units) and 1 parīsu’.

Figure 7: TODO [Cha12, p. 61]

This is particularly true of the signs ◀ , $\text{◀}\text{◀}$, $\text{◀}\text{◀}\text{◀}$ and $\text{◀}\text{◀}\text{◀}\text{◀}$, whose form explicitly denotes the fractions 1/6, 2/6, 3/6, and 4/6 of the barig capacity measure written ◀ in Mesopotamia—also transcribed by Assyriologists as 1 bán, 2 bán, 3 bán, and 4 bán with reference to the bán measure worth 1/6 of the barig. At Ebla, the sign ◀ is most often associated with the *parīsu* measure, while the signs ◀ , $\text{◀}\text{◀}$, $\text{◀}\text{◀}\text{◀}$ and $\text{◀}\text{◀}\text{◀}\text{◀}$ refer to 1, 2, 3,

Figure 8: TODO [Cha12, p. 64]

shape. The principle of notation is additive: each sign is noted as many times as necessary (e.g., $\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}\text{◀}$ transliterated as 2(šar₂) 1(geš’u) 3(u), means $2 \times 3600 + 1 \times 600 + 3 \times 10$). The system is based on an alternation of factors ten and

Figure 9: TODO

might think of one fabric and a half,¹¹ but the presence of notations with “2▷2▽”, “3▷3▽”, and “6▷6▽” (Fig. 1) elements excludes that one deals with fractions, as these notations are not consistent with those of Šuruppag’s weight measurement system.¹² The notation “1▷gada” in o. ii 1 and r. vi 1, along with the total of “39



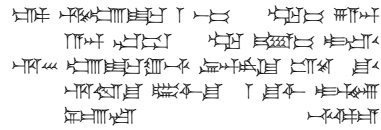
Fig. 1. Combinations of numerals attested in Š. 742.

Figure 10: Discussion of the contrast between ▷ and ▽ numerals in [Gor23, p. 162].

as, for example, in TM.75.G.3125 = ARET III 107 o. iv 1, “4▷ ‘a₃-da-um^{u9}-2 4▽
aktum 4▷ib₂^{u9}×3▽ sa₆ gunu₃” (Fig. 2).

Figure 11: Transliteration in [Gor23, p. 163] of [P242293, recto 4, 1] incorporating untransliterated numerals.

3.5 Non-numeric usage



The beginning of the scribal art is a single wedge. That one has six pronunciations; it also stands for ‘sixty’⁵¹. Do you know its reading⁵²?

Examenstext A

Many of the cuneiform numerals are used with a logographic or phonetic value. For example, the sign — has, *inter alia*, the values aš, rum, and dili. While the horizontal numerals are most frequently written ▷ in the Early Dynastic period⁵³, such non-numeric usage is almost⁵⁴ always written —, for instance:

- in personal names in administrative texts, such as the following, which all contain ▷ numerals:
 - — in [P010424; P010458; P010459] from ED IIIa أبو صلابيخ,
 - — in [P010960] from ED IIIa Šuruppag,
 - — in [P251641] from ED IIIb Adab,

encoded $ib_2 \times 3! = \text{𐎶} \times \text{𐎶} = \text{𐎶𐎶}$.

⁴⁹TODO cite the EbDA one.











⁵⁰TODO(egg): On the order cite TSS 188, Friberg2007 p. 148 and any of the usual suspects on the haphazard order of signs in early texts; contrast P274845, P241764.

⁵¹The reader will recall that $\eta e\dot{s}_2$ is written 𐎶, with a larger wedge than 𐎶; however, these signs have merged by the time Examenstext A is composed.

⁵²Besides $\eta e\dot{s}_2$, a look at [VT+14] shows that the values diš, ge₃, makkaš, sanjak₄, and tal₄ are attested both in [TJV17] and in lexical lists. The sign is also used for the Akkadian word *ana* in the Neo-Assyrian period.

⁵³A CDLI search for “(asz@c)” finds 3296 ED texts, while a search for “(asz)” finds 81 ED texts, of which 46 also contain “(asz@c)”.





⁵⁴Exceptions are discussed in section 3.7.1.

-    in [P252866] from ED IIIb Adab,
-    in [P298637] from ED IIIb Umma;
- in the Sumerian word  $\text{u}_2\text{-rum}$, “property” in ED IIIb Nirsu administrative texts which contain \triangleright numerals, such as [P020006; P020008; P020018; P020024; P020030];
- in lexical texts:
 - in the divine name    $\text{— } \text{L}$ in the lexical texts [P010570; P010572], where the entries are prefixed with \triangleright .
 - in the word $\text{— } \text{dili}$, “small fish” in [P010578], witness to Early Dynastic Fish,
 - in the same word with a determinative, $\text{— } \text{dili}^{\text{ku}_6}$, in [P010586], witness to Early Dynastic Food, which starts with \triangleright numerals.

This is a clear contrast between — and \triangleright in this period, and genuine ambiguity can arise if it is lost; for instance, the personal name $\text{— } \text{—}$ occurs on its own line in the aforementioned administrative texts; a line $\triangleright \text{—}$ would instead be read as “one slave”.

3.6 Limited benefits of diachronic encoding for numerals

The argument in favour of diachronic encoding is that it facilitates interoperability in a variety of use cases, as we have outlined in section 3.1. While these benefits are real and now visible for cuneiform signs, similar considerations are not generally applicable to curviform numerals.

Diachronic reference works such as sign lists and dictionaries tend to not include numbers, or when they do, they treat them separately, and include signs such as — that have both numeric and non-numeric values in both the main list and the section on numbers. For instance, [Sch35, pp. 123 sqq.] lists all of —  together with \triangleright , while — , — , and — , and only those, appear at the beginning of the sign list, since they have non-numeric values⁵⁵. [Cat13, p. 58] has the numeric signs \triangleright , — , — , whereas non-numeric — is at the beginning of the sign list, where its values *aš* and *rum* are listed. For signs with both non-numeric and numeric usage, [Dei22] writes *s. die Zahlz.* throughout the main list; LAK 1 — thus reappears at LAK 829 together with \triangleright , — , and — . One should note [Bor10], which has numbers throughout the sign list; but that sign list does not show glyphs predating the Old Babylonian period, nor does it comprehensively cover the numerals used in the Ur III and Old Babylonian periods, as, for instance, it does not have  —  used in system $G_{\text{Ur III/OB}}$.

Composite texts rarely have witnesses both from the Early Dynastic period and later; the kinds of texts that do, chiefly lexical and literary texts, do not contain numbers to the extent that administrative texts do. Further, there tend to be changes⁵⁶ to the text between Early Dynastic and later witnesses that prevent a diachronic encoding of such composites. For numerals, the switch from \triangleright to — numerals prevents diachronic encoding even if \triangleright were unified with — . For instance, the lexical list Early Dynastic Food, already mentioned in section 3.5, contains some numbers, and has a witness from the Old Akkadian period covering these numbers: [P215653, a 1’–6’]; however, they are written with — numerals, whereas they are writ-

⁵⁵Non-numeric values of — were discussed in section 3.5; — has the values *man*₃ and *min*₅, and is used for the word *didli*, “several, various”; — has the value *eš*₆.

⁵⁶TODO comment on the ED witnesses to the instructions of Šuruppak

ten with \triangleright numerals in the Early Dynastic witnesses; since \uparrow and \leftarrow are distinct⁵⁷ characters, the $\triangleright \leftarrow$ unification does not help.

More generally, since numbers are so deeply tied to metrology, and since metrological systems change between the Early Dynastic and later periods⁵⁸, there is little opportunity for a diachronic representation of numeric quantities.

In the case of analyses such as [Romach2023], it is interesting to note that numeric expressions are removed prior to the conversion of the corpus to Unicode cuneiform for further analysis.

3.7 Compatibility considerations

A disunification twenty years after the fact, affecting all numerals, would ordinarily be a serious compatibility issue. Fortunately, with the exception of one character discussed below, we are not aware of any font using curviform glyphs for the already-encoded numerals. In fact we are not aware of any font designed for a style earlier than Old Babylonian, except for fonts mimicking the representative glyphs from the code charts, which are primarily Ur III, but sometimes earlier or later, as described in [UTR56, §2.4]. The lack of dedicated Ur III fonts may be explainable by the chart-like fonts⁵⁹ being good enough for most purposes; the lack of Early Dynastic fonts, by the aforementioned issues with numeral unification making the representation of any text with numerals intractable.

3.7.1 The case of ŠAR₂

The character U+122B9 \bullet CUNEIFORM SIGN ŠAR2 has a circular reference glyph.

In most texts from the Early Dynastic IIIb and Old Akkadian period⁶⁰, a contrast between non-numeric šar₂ written \diamond and numeric 1(šar₂) written \bullet can be observed, similar to the contrast between \leftarrow and \triangleright previously discussed in section 3.5. However, in lexical lists from Šuruppak and Ebla⁶¹, as well as in the *Stèle des vautours*, non-numeric šar₂ is curviform:

- $\ast \text{𒌦} \text{𒌦} \text{𒌦} \bullet$ and $\ast \text{𒌦} \bullet \text{𒌦} \text{𒌦}$ in [P010566];
- $\bullet \text{𒌦} \text{𒌦}$ and $\ast \bullet \text{𒌦} \text{𒌦}$ in [P010576];
- $\bullet \text{𒌦}$ in [P240986]⁶²;
- $\bullet \text{𒌦} \text{𒌦}$ in [P222399, obv. 17, 9, 18, 11, 22, 12]⁶³.

⁵⁷Besides the contrasts in numeric usage mentioned in section 3.3.3, these characters are clearly not unifiable because of the many contrasts in non-numeric usage between them; several values of \leftarrow which are not shared with \uparrow have already been mentioned, but perhaps most striking is the fact that, in the Neo-Assyrian period, \leftarrow is used for the preposition *ina*, “in”, and \uparrow for the preposition *ana*, “to”.

⁵⁸TODO cite a few things here.

⁵⁹Most prominently Noto Sans Cuneiform, a system font on both Windows—as part of Segoe UI Historic—and macOS.







⁶⁰For example, in personal names:


- $\text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$ in [P020019] from ED IIIb Nirsu;
- $\text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$ in [P020182], also from ED IIIb Nirsu;
- $\text{𒌦} \text{𒌦}$ in [P222186] from ED IIIb Umma;
- $\text{𒌦} \text{𒌦} \text{𒌦}$ in [P235312] from Old Akkadian Umma.





⁶¹TODO Mention other ways in which these are archaizing

⁶²From copy in [ELLES].




⁶³Note however $\ast \text{𒌦} \text{𒌦} \text{𒌦}$ on [P222399, obv. 6, 17]. Curviform non-numeric šar₂ is clearly archaizing in ED IIIb Nirsu; one might suppose that the scribe slipped into their modern ways here. TODO add a photo.





It *would* be disruptive to the diachronic representation of text if non-numeric šar₂ were to have two different representations. The character U+122B9 CUNEIFORM SIGN SHAR2 should therefore be used in those cases, with its curviform glyph , identical to the glyph of the proposed U+12579  CUNEIFORM NUMERIC SIGN ONE N45. Since the archaizing style of texts wherein non-numeric šar₂ is curviform solidly predates the transition from  to  in the relevant metrological systems, there is no need to represent a - contrast, so these characters can have the same glyph in specialist archaizing Early Dynastic fonts.










Since cuneiform U+122B9 CUNEIFORM SIGN SHAR2 effectively merges with U+1212D  CUNEIFORM SIGN HI, the reference glyph should remain as it is, *i.e.*, curviform, so that the contrast between reference glyphs within the Cuneiform block remains clear; see [UTR56, §2.4]. Since system fonts follow the reference glyphs, and since extant specialist fonts target styles where U+122B9 is unambiguously cuneiform, there are no compatibility issues.

Note that in rare cases, such as [P222243] from ED IIIa Adab, non-numeric  (here with the value rum) is written . It is out of scope for this proposal to decide whether such occurrences should be treated as anomalous spellings, encoded as U+12550  cuneiform numeric sign one N01, or as stylistic distinctions, encoded as U+12038 CUNEIFORM SIGN ASH with a curviform glyph. In practice this would often be determined by the transliteration from which the cuneiform text is generated; it is noteworthy that as of this writing, the CDLI transliteration (UR2-1(aš@c)) and the ePSD2 one (uru₈^{rum}) of this word disagree on that aspect. Since  has a cuneiform reference glyph, this does not pose any compatibility concerns.

3.7.2 Transliteration

An important feature of the encoding is that, in order to support input and bulk conversion of transliterated corpora to Unicode cuneiform, it should not represent distinctions that are finer than those recorded in typical transliterations; thus, while some older forms of BIL₂ can be described as  NE×KASKAL or  NE×PAP⁶⁴, they are typically all transliterated bil₂, and therefore are all represented by the character U+1224B  CUNEIFORM SIGN NE SHESHIG, its name notwithstanding, as described in [UTR56, §2.5].

The situation is more complicated for numbers. Many transliterations do not represent the type of numeral used, instead interpreting the whole numeric expression and transcribing it with delimiters or units as needed to disambiguate. For instance,  from [P305639] may be transliterated as 95 gur, as in [Feu04, vol. 2, p. 62]. The numerals may also be transliterated separately, but solely by their values in terms of the overt unit, as in EbDA transliterations: the aforementioned  from [P240533, recto 3, 3] is transliterated “20-1-1/2 gu₂-bar 7 nig₂-sagšu 2-1/2 an-zam_x⁶⁵ za”, reading both  and  as 1/2, but not distinguishing them.

In particular, these transliterations do not differentiate between  and  numerals, nor between  and  numerals. For instance, the aforementioned  from [P242293, recto 4, 1] is transliterated “4 'a₃-da-um^{tug₂}-II 4 aktum^{tug₂} 4 ib₂-III gun₃ sa₆^{tug₂}” in EbDA, with no distinction between the  and . Since  and  numerals are separately encoded, the numeric ex-

⁶⁴As on [P249253].

⁶⁵As of this writing, EbDA actually has an-zam_x, with U+1D6A GREEK SUBSCRIPT SMALL LETTER CHI.

pressions in such transliterations cannot be transformed into Unicode cuneiform without additional context, regardless of curviform–cuneiform unification.

In metrological systems such as systems $G_{Ur III/OB}$ and $C_{Ur III/OB}$ where some units are indicated by the type of numeral rather than an overt unit sign, it is common practice to add the unit in parentheses in transliteration; for instance, $\text{𒀭} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$ from [P386847] is transliterated “1(eše₃) 5½ iku⁶⁶ 7 sar” in [Feu04, vol. 2, p. 176], and $\text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$ from [P307255] is transliterated “1(n⁶⁷) 2(b) 7 ½ sila₃” in [Feu04, vol. 2, p. 151].

This practice has been generalized to systematically indicate numeral shape; this is in particular the case in CDLI, where the transliterations of some of the above examples are “1(gesz2) 3(u) 5(asz) gur” for $\text{𒀭} \text{𒌦} \text{𒌦} \text{𒌦}$, “1(esze3) 5(iku) 1/2(iku) GAN2 7(disz) sar” for $\text{𒀭} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$, and “3(barig) 2(ban2) 7(disz) 1/2(disz) sila3” for $\text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$. CDLI and ePSD2 both distinguish curviform from cuneiform numerals in transliteration: the length $\text{𒀭} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦} \text{𒌦}$ from [P020129, rev. 2, 1] is transliterated “6(gesz2@c) 3(u@c) {ninda}nindax(DU) 1/2(asz@c) 4(disz@t) gi” in CDLI, and “6(geš₂^c) 3(u^c) ninda ninda_x(DU) 1/2(aš^c) 4(diš^t) gi” in ePSD2. Another example is [Molina2014], which uses *1a* for 𒀭 , *1d* for 𒌦 , *1ac* for 𒌦 , *1dc* or $\frac{1}{2}dc$ for 𒌦 depending on reading, etc. The literature on the Uruk and Early Dynastic I–II periods uses a different set of transliteration conventions that also disambiguate numeral shapes, as will be discussed in section 4.

While there exist transliterations that distinguish 𒀭 from 𒌦 but not 𒀭 from 𒀭 , such as the ones used in [DCCMT], the trend, especially in more recent works in third millennium studies, seems to be to represent numeral shape; for example, [Maiocchi2024] gave an example of the input syntax used by the new “Urban Economy Begins” project as “10 + 5c(GUR) + 2(BARIGA) + 1(BAN2)” for $\text{𒀭} \text{𒀭} \text{𒀭} \text{𒀭} \text{𒀭}$, with a *c* indicating that the GUR numerals are curviform, and the parenthetical GUR indicating that these are 𒀭 rather than 𒌦 numerals.

3.8 Conclusions

Co-occurrences of curviform and cuneiform numerals are not anecdotal in the Early Dynastic period. Instead, they represent contrasts between metrological systems, between individual units within metrological system, and between numeric usage and phonetic or logographic usage. This contrastive usage is reflected in modern publications.

While it would be technically possible to handle this contrast as a stylistic distinction, this approach has no real benefit, and is highly inconvenient, as it requires single numeric expressions to systematically use multiple fonts. Further, if that contrast is lost in plain-text interchange, the text can be misinterpreted: 𒀭 is a length of three ropes, but 𒀭 is an area of three bur₃; 𒀭 could be read as one 𒀭 and one 𒌦 , where 𒀭 would be one and a half 𒀭 ; 𒀭 is a personal name, but 𒀭 would be “one slave”.

At the same time, contrary to most disunifications, the separate encoding of curviform numerals poses no serious compatibility issues for existing fonts or encoded corpora, nor does it, in general, introduce new issues with transliterated third millennium corpora. The oddity of 𒀭 requires some explanation, but does not

⁶⁶TODO say something about this reading

⁶⁷TODO comment on nigida.

pose any architectural issues, and is not fundamentally different from the other mergers and splits encountered in the cuneiform script.

4 Rationale for ED–Uruk numeral unification

TODO mention the bariga silliness in the CDLI transliteration of Gori’s paper.

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

N13 not attested in CDLI (N_{17} , $12N_{14}$, etc.) 7(diš *tenû*)

6.2 Stacking patterns

The already-encoded numerals in the Cuneiform Numbers and Punctuation block distinguish some stacking patterns; for instance 9I is encoded both as U+12446 𐎶 and as U+1240E 𐎶𐎶. This is in part due to contrastive usage of stacking patterns; besides I and 𐎶 which are characteristic of bariga measures, four bariga is written 𐎶 even where 4I is written 𐎶, as in [P255010; P292843]. This is also for compatibility with distinctions made in reference works and in some non-numeric transliterations; for instance, 𐎶 is [Bor10, No. 860] and has the value limmu, and 𐎶 is [Bor10, No. 852] and has the value limmu₅.

However, there is no such practice when it comes to numeric transliterations⁶⁸. The stacking patterns used in numeric expressions change over time, and do not carry any semantics; the typical 4(diš) is 𐎶 in Ur III, but 𐎶 in the Neo-Assyrian period.

Since they are not listed separately in classical sign lists nor assigned any different values, the stacking patterns from earlier periods are not separately encoded; for instance, in ED IIIb Nirsu, 𐎶 2(u) often has one 𐎶 atop another. There is no evidence of a need to encode such variant stacking patterns; these distinctions would be incompatible with the state of the art in numeric transliterations, and are not needed to represent reference works. Instead, they should be considered style variants, and an ED IIIb Nirsu font should have an appropriate glyph for U+12399 𐎶 CUNEIFORM SIGN U U.

Likewise, many stacking patterns are attested for the curviform numerals proposed in this document, and it is not proposed to separately encode them. Idiosyncratic stacking patterns are in fact particularly common in Early Dynastic and earlier tablets, as they are structured in rectangular cases rather than lines, so that

⁶⁸The Sumerian word limmu means “four”, so limmu and limmu₅ are still numbers. The distinction here is between usage in transliterations of phrases such as 𐎶𐎶𐎶𐎶 𐎶𐎶𐎶𐎶 lugal an-ub-da limmu₅-ba-ke₄ (king of the four quarters) or of names, and of numeric expressions such as 𐎶𐎶 4(diš) sila₃.



Figure 12: The layout of case [P020066]; the numeral ●● is spread across two lines. The text is read in the order ●●▷▷▷▷ ↗ ↘, “twenty-two oxen, one year old”.

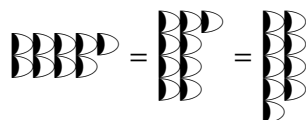


Figure 13: Three stacking patterns for U+12573 CUNEIFORM NUMERIC SIGN NINE N34. The one on the left is the reference glyph, used in Uruk III [P003499; P004430], and widely afterwards, *e.g.*, ED IIIa Šuruppag [P010678], ED IIIb Nirsu [P020057], Old Akkadian Umma [P212464]. The ones in the middle and right are used in two Uruk IV tablets [P001243; P004500]. All the Uruk examples are transliterated 9(N34) in CDLI, and the later ones 9(gesz2@c).

numerals may be laid out across the case in whichever way fits the available space; this is illustrated in Figure 12. Note that in that figure, the numerals need to be considerably enlarged in order to reproduce the layout of the tablet: ▷ has the same height as ▷▷. This is impractical when these numerals are set in text that contrasts them with the larger ▷▷, and inconsistent with actual practice when typesetting these numerals, as illustrated in section 3.4.5: reproducing the layout of tablets is not within the scope of plain text.

The reference glyphs use stacking patterns that are common in the Early Dynastic period, but that are also attested in the Uruk period; the Uruk period also frequently features numerals that use a more vertical layout, as illustrated in Figure 13. The later, more horizontal styles were chosen for two reasons: for the numerals used in the third and fourth millenium, usage in third millenium scholarship will be more frequent; and the horizontal layout poses fewer layout difficulties when set in lines of non-cuneiform text, as most modern scholarship is. Indeed, the absolute size of the indents ▷, ▷▷, ●, and ●● must remain consistent across the numeral series, lest a ▷▷ numeral be confused with an ▷ numeral. Since the single indents are frequently used in running text, as illustrated in section 3.4.5, they need to be large enough that the vertical stacking patterns are impractical.

Variant stacking patterns, if needed, may be handled at a higher level as stylistic distinctions; Figure 13 uses OpenType stylistic alternates, and Figure 12 rotates the character ●●, in both cases preserving the plain text backing.

Acknowledgements

TODO(egg): Something about the Vanséveren fonts

References

Artefacts

- [P020054] VAT 4731. [För16, 40 p.14]. Vorderasiatisches Museum.
CDLI: [P020054](#).
- [P020129] VAT 04713. Vorderasiatisches Museum.
CDLI: [P020129](#).
ORACC: [epsd2/corpus/P020129](#).
- [P102305] X.3.139. Michael C. Carlos Museum, Emory University.
CDLI: [P102305](#).
- [P215653] AS 15375 21. Musée du Louvre.
CDLI: [P215653](#).
ORACC: [dcclt/corpus/P215653](#).
Louvre Collections: [ark:/53355/cl010436723](#).
- [P221266] AO 13825. Musée du Louvre.
CDLI: [P221266](#).
ORACC: [epsd2/corpus/P221266](#).
Louvre Collections: [ark:/53355/cl010138527](#).
- [P221291] AO 13850. Musée du Louvre.
CDLI: [P221291](#).
ORACC: [epsd2/corpus/P221291](#).
- [P221305] AO 13864. Musée du Louvre.
CDLI: [P221305](#).
ORACC: [epsd2/corpus/P221305](#).
- [P222399] *Stèle des vautours*. AO 50; AO 2346; AO 2347; AO 2348; AO 16109.
Musée du Louvre.
CDLI: [P222399](#).
- [P232278] *Gudea E*. AO 6. Musée du Louvre.
CDLI: [P232278](#).
ORACC: [etcsri/Q001544](#).
- [P232280] *Gudea G*. AO 7. Musée du Louvre.
CDLI: [P232280](#).
ORACC: [etcsri/Q001546](#).
- [P240531] TM.75.G.00265. Idlib, Syria: National Museum of Syria.
CDLI: [P240531](#).
EbDA: [1415](#).
- [P240532] TM.75.G.00266. Idlib, Syria: National Museum of Syria.
CDLI: [P240532](#).
EbDA: [1324](#).
- [P240533] TM.75.G.00267. Idlib, Syria: National Museum of Syria.
CDLI: [P240533](#).
EbDA: [1379](#).
- [P240548] TM.75.G.00302. Idlib, Syria: National Museum of Syria.
CDLI: [P240548](#).
EbDA: [1350](#).

- [P240579] TM.75.G.00341. Idlib, Syria: National Museum of Syria.
CDLI: [P240579](#).
EbDA: [1364](#).
- [P240609] TM.75.G.00440. Idlib, Syria: National Museum of Syria.
CDLI: [P240609](#).
EbDA: [1378](#).
- [P240653] TM.75.G.00535. Idlib, Syria: National Museum of Syria.
CDLI: [P240653](#).
EbDA: [1382](#).
- [P240654] TM.75.G.00536. Idlib, Syria: National Museum of Syria.
CDLI: [P240654](#).
EbDA: [1383](#).
- [P240655] TM.75.G.00537. Idlib, Syria: National Museum of Syria.
CDLI: [P240655](#).
EbDA: [1358](#).
- [P240675] TM.75.G.00557. Idlib, Syria: National Museum of Syria.
CDLI: [P240675](#).
EbDA: [1371](#).
- [P240697] TM.75.G.00579. Idlib, Syria: National Museum of Syria.
CDLI: [P240697](#).
EbDA: [1381](#).
- [P241708] TM.75.G.02143. Idlib, Syria: National Museum of Syria.
CDLI: [P241708](#).
EbDA: [3173](#).
- [P241904] TM.75.G.02346. [[Arc89](#), p. 6]. Idlib, Syria: National Museum of Syria.
CDLI: [P241904](#).
EbDA: [3183](#).
- [P242293] TM.75.G.03125. Idlib, Syria: National Museum of Syria.
CDLI: [P242293](#).
EbDA: [217](#).
- [P249253] *Code de Hammurabi*. Sb 8. Musée du Louvre.
CDLI: [P249253](#).

Unicode documents

- [L2/03-162] M. Everson and K. Feuerherm. *Basic principles for the encoding of Sumero-Akkadian Cuneiform*. 25th May 2003.
UTC: [L2/03-162](#).
- [L2/03-393R] M. Everson, K. Feuerherm and S. Tinney. *Preliminary proposal to encode the Cuneiform script in the SMP of the UCS*. 3rd Nov. 2003.
UTC: [L2/03-393R](#).
- [L2/04-036] M. Everson, K. Feuerherm and S. Tinney. *Revised proposal to encode the Cuneiform script in the SMP of the UCS*. 29th Jan. 2004.
UTC: [L2/04-036](#).
- [L2/04-099] L. Anderson. *Unification of Cuneiform Numbers*. 2004.
UTC: [L2/04-099](#).

- [L2/04-189] M. Everson, K. Feuerherm and S. Tinney. *Final proposal to encode the Cuneiform script in the SMP of the UCS*. 8th June 2004.
UTC: [L2/04-189](#).
- [L2/24-159] P. Constable, ed. *Minutes of UTC Meeting 180* (23rd–25th July 2024). 29th July 2024.
UTC: [L2/24-159](#).
- [Uni16] The Unicode Consortium. *The Unicode Standard*. Version 16.0.0. The Unicode Consortium, 10th Sept. 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/>.

Other documents

- [Arc15] A. Archi. *Ebla and Its Archives. Texts, History, and Society*. Studies in ancient Near Eastern records 7. Walter de Gruyter, 2015.
ISBN: 978-1-61451-716-0.
DOI: [10.1515/9781614517887](https://doi.org/10.1515/9781614517887).
- [Arc89] A. Archi. “Tables de comptes eblaïtes”. In: *Revue d’assyriologie et d’archéologie orientale* 83.1 (1989). Ed. by P. Amiet and P. Garelli, pp. 1–6. ISSN: 0373-6032.
- [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.
- [Cat13] A. Catagnoli. *La paleografia dei testi dell’amministrazione e della cancelleria di Ebla*. Quaderni di Semitistica 9. Università di Firenze, 2013.
ISBN: 8890134054.
- [Chao3] G. Chambon. “Archaic Metrological Systems from Ur”. In: *Cuneiform Digital Library Journal* 2003.5 (23rd Dec. 2003). ISSN: 1540-8779.
eprint: http://cdli.ucla.edu/pubs/cdlj/2003/cdlj2003_005.html.
- [Cha12] G. Chambon. “Numeracy and Metrology”. In: *The Oxford Handbook of Cuneiform Culture*. Ed. by K. Radner and E. Robson. Oxford University Press, 18th Sept. 2012, pp. 51–67.
ISBN: 9780199557301.
DOI: [10.1093/oxfordhb/9780199557301.013.0003](https://doi.org/10.1093/oxfordhb/9780199557301.013.0003).
- [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.

- [Dei22] A. Deimel. *Liste der archaischen Keilschriftzeichen von Fara*. Wissenschaftliche Veröffentlichungen der Deutschen Orient-Gesellschaft 40. J. C. Hinrichs'sche Buchhandlung, 1922.
- [Eng98] R. K. Englund. "Texts from the Late Uruk Period". In: *Mesopotamien. Späturuk-Zeit und Frühdynastische Zeit*. Orbis Biblicus et Orientalis 160/1. 1998, pp. 13–233. ISBN: 3-7278-1166-8.
- [Feu04] K. G. Feuerherm. "Abum-waqar and His Circle. A Prosopographical Study". PhD thesis. University of Toronto, 2004.
- [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*. 12th Apr. 2016. eprint: https://cdli.ox.ac.uk/wiki/doku.php?id=ed_iii_metrological_systems.
- [Gor23] F. Gori. "On Lapis Lazuli and Linen in Šuruppak Texts. An Analysis Through the Lens of Ebla Studies". In: *Studia Eblaitica* 9 (2023), pp. 160–166. ISSN: 2364-7124.
- [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. Jauhiainen and H. Jauhiainen. "Advancing Cuneiform Text Dating Through Automatic Analysis". 69th Rencontre Assyriologique Internationale (8th–12th July 2024). 11th July 2024 14:00.
- [Kre98] M. Krebern timer. "Die Texte aus Fāra und Tell Abū Šalābīḥ". In: *Mesopotamien. Späturuk-Zeit und Frühdynastische Zeit*. Orbis Biblicus et Orientalis 160/1. 1998, pp. 235–427. ISBN: 3-7278-1166-8.
- [Lec12] C. Lecompte. "Des chiffres et des digues: à propos de deux textes présargoniques de Ġirsu et d'une notation numérique inhabituelle". In: *Altorientalische Forschungen* 39.1 (Dec. 2012), pp. 81–86. DOI: [10.1524/aof.2012.0006](https://doi.org/10.1524/aof.2012.0006).

- [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.
- [Mil+07] L. Milano, M. Maiocchi, F. Di Filippo, R. Orsini, E. Scarpa, M. Surdi et al., eds. *Ebla Digital Archives*. 2007–. eprint: <http://ebda.cnr.it/>.
- [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 Ḫammu-rāpi*. dubsar 4. Zaphon, 2022.
- [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*. 18th Dec. 2019. ORACC: [doc/help/editinginatf/metrology/metrologicaltables](http://oracc.org/doc/help/editinginatf/metrology/metrologicaltables).
- [Rob22] E. Robson. “Overview of Metrological Systems”. In: *The Digital Corpus of Cuneiform Mathematical Texts*. 2022. ORACC: [dccmt/Metrology](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 (8th–12th July 2024). 10th July 2024 12:00.
- [Sch10] W. Schramm. *Akkadische Logogramme*. Göttinger Beiträge zum Alten Orient 5. Universitätsverlag Göttingen, 2010. ISBN: 978-3-941875-65-4. DOI: [10.17875/gup2010-511](https://doi.org/10.17875/gup2010-511).
- [Sch35] N. Schneider. *Die Keilschriftzeichen der Wirtschaftsurkunden von Ur III*. Editrice Pontificio Istituto Biblico, 1935.
- [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>.
- [VT+14] N. Veldhuis, S. Tinney et al., eds. *Oracc Sign List*. 2014–. eprint: <http://oracc.org/osl/>.