Archaic cuneiform numbers

Robin Leroy, Anshuman Pandey, and Steve Tinney

2024-09-02

Contents

1 Summary

This document proposes encoding some numerals used in the Uruk and Early Dynastic periods in conjunction with the Sumero-Akkadian cuneiform script[1] and the proto-cuneiform script[2]. 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", numerals[3] should however *not* be unified with the already-encoded cuneiform numerals[4]. 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.]

---

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

[2]ISO 15924: Pcun, not yet encoded.

[3]Impressed 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.*, . The terms "curved", "curviform", and "round" can be found in the literature. We avoid the term "round" here as it has other meanings in the context of numbers. We use "curviform" in this document as, being the least common term, it is least likely to lead to confusion, and "CURVED" in the character names for consistency with documentation about the modifier @c used in machine readable ATF transliterations [**inlineATF**].

[4]Impressed into clay using a stylus with a trihedral end: 𒀸 (stylus held horizontally), 𒁹 (vertically), (diagonally) 𒌋 (with the head of the stylus), (stylus pressed deeper, forming a larger wedge), (combining and 𒌋), etc.

<table>
<tr><td></td><td>Uruk III & earlier</td><td>ED – Ur III</td><td>OB & later</td></tr>
<tr><td rowspan="2">Numerals</td><td colspan="2">This proposal</td><td></td></tr>
<tr><td></td><td colspan="2" rowspan="2">Existing Xsux</td></tr>
<tr><td>Non-numeric signs</td><td>Future Pcun</td></tr>
</table>

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

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 𒅎 in texts from Early Dynastic IIIa Šuruppag as in the character code charts, 𒅎 later in the third millenium[5], 𒅎 in Old Babylonian cursive, 𒅎 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 lists[6] and dictionaries[7], and of composite texts[8]. 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 applications[10].

[5] Merging with U+1224E cuneiform sign ni2.

[6] Notably [OSL] and the online edition of [MZL] in [eBL, Signs].

[7] Notably [ePSD2] and the online edition of [Sch10] in [eBL, Dictionary].

[8] For example, there are Neo-Assyrian and Neo-Babylonian copies parts of the laws of 𒄩𒄠𒈬𒊏𒁉, as well as Old Babylonian copies in both archaizing and cursive styles. Because of damage on the stele [P249253], some sections are known only from those copies. See [Oel22, pp. 110 sqq.].

[9] Attendees may recall the summary given on the third day of UTC #180, as recorded in [L2/24-159]. Other readers may refer to [Svä+24, pp. 242, 148].

[10] For instance, Old Babylonian grammar may be taught in the Neo-Assyrian script, as in [Cap02].

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 encoding[11] 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 Ŋirsu. The excerpt cited, lines 1–3 of column 1 of the obverse, is as follows:

[12]						
1(ŋeš$_2$)	1(u)	1/2(diš)	5(diš *tenû*)	gi	us$_2$	sa$_2$
7.5 (ropes)			5	reed	side	equal

[13]				
3(u)	6(diš *tenû*)	gi	saŋ	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)
field-this				

tug$_x$(LAK483)-si-ga-kam		
tugsiga	=ak	=am
ploughed	=GEN	=COP

The argument made in [L2/04-099, p. 4] is that this is comparable to a stylistic distinction such as[14]

> 465 metres, equal lengths
> 198 metres, equal widths
> this field is 9, 18 hectares of ploughed land

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

[11] At that time scoped to the répertoire of the Ur III period and later, see [L2/03-162, p. 1], although many disunifications, such as ≠ , were informed by Early Dynastic distinctions.

[12] As 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.

[13] Note 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 .

[14] We 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.

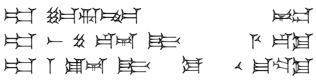
curviform numerals should be seen as a formatting distinction, rather than one that should be represented in plain text, and insists that the encoding should capture the lineal historical descent of those signs, presumably to take advantage of the benefits of diachronic encoding described in section 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 𒁹/𒁹 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 A primer on classic Ur III and Old Babylonian metrologies



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

Edubba'a D

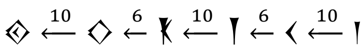
Before diving into the usage of the curviform numerals in the Early Dynastic period to explain the constrast with cuneiform numerals, it is useful to understand the usage of the already-encoded characters in the Ur III and Old Babylonian periods.

As is well known[15] a sexagesimal place value system (SPVS) was used in Mesopotamia from the late third millenium onwards. One should bear in mind, however, that other systems were used; the SPVS was primarily used in calculations, with results being expressed in non-positional systems [Rob08, p. 76; Rob22]. The digits 1–59 of the SPVS have inner structure which is reflected in the encoding: the digits 1–9 are the individual characters 𒁹–𒐈, the multiples of ten (10–50) are 𒌋–𒐐, but the other digits 11–59 are sequences 𒌋𒁹–𒐐𒐈; in effect the base-sixty digits are themselves written in base ten, with a different set of symbols for the tens place. This reflects the origin of the sexagesimal place value system; it derives from a *non-positional* system, hereafter the *cuneiform discrete counting system* $S_{\text{Ur III/OB}}$, which had different signs for the units 𒁹–𒐈, tens 𒌋–𒐐, sixties 𒐕–𒐝 (with larger wedges than the units), multiples of six hundred 𒐞–𒐦, multiples of three thousand six hundreds 𒊹–𒐮, and multiples of thirty-six thousand 𒐯–𒐳.

[15] See, *e.g.*, [Uni16, §22.3.3, *sub* "Cuneiform Numerals"].

3.3.1 The discrete counting system

The relations between the values of the signs in the cuneiform discrete counting system may be summarized by the following factor diagram[16], where the number over arrow indicates the multiple of the preceding sign (right of the arrow) corresponding to the following sign (left).

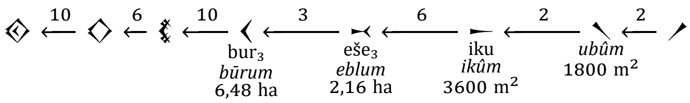


$(S_{\text{Ur III/OB}})$

For example, the number $1729 = ((2 \times 10 + 8) \times 6 + 4) \times 10 + 9 = 28 \times 60 + 49$ would be written 𒐙𒐜𒐏𒐝 in the discrete counting system, and 𒎙𒐜𒐏𒐝 in the sexagesimal place value system.

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]:

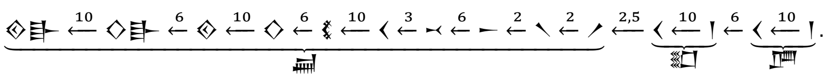


$(G_{\text{Ur III/OB}})$

Note that for the range of areas given above[17], this system does not use any symbols separate from the numerals for the individual units (*ubûm*, *ikûm*, *eblum*, and *būrum*). As mentioned in [Rob19], the whole numeric expression for the area would be followed by the sign 𒃷 functioning as punctuation[18], but the numerals are tied to the metrology; thus a surface of 5 *būr* 1 *ebel* 4 *ikû* (100 *ikû*, 36 ha) would be written[19] 𒐊𒑏𒐂𒃷. 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 𒊹 is equal to sixty times 𒌋 in the area system, but to three hundred and sixty times 𒌋 in the discrete counting system.

[16] These diagrams, which have become standard in discussions of Mesopotamian metrology, originate with [Fri78, p. 10], where they are called *step-diagrams*.

[17] For areas smaller than a quarter *ikûm*, an overt unit is used, with 1 *mūšarum* (36 m²) written 𒁹𒊬, equal to one hundredth of an *ikûm*, then sexigesimally subdivided in 60 𒂆 (shekels). For areas greater than 3600 *būrū*, the 𒊹 and 𒊹 numerals are reused with a suffix 𒃲 (gal, Sumerian: big), as follows [Rob08, p. 295 n. b and c; Fri07, p. 378; Rob19]:



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

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

3.3.3 The capacity system

Another such system of note is the one for capacities[20] [Fri07, p. 376; Rob19],

𒐬 ←10– 𒊹 ←6– 𒐞 ←10– 𒁹 ←6– 𒌋 ←10– 𒀸 ←5– 𒁹 ←6– 𒑏 ←10– 𒁹𒋡, $(C_{\text{Ur III/OB}})$

gur *kurrum* (𒐬 … 𒀸: 𒄥); bariga *parsiktum*; ban$_2$ *sūtum*; sila$_3$ *qûm* 1 l

where the numerals for ban$_2$ are 𒑏, 𒑐, 𒑑, 𒑒, and 𒑓, and those for bariga are 𒁹, 𒑔, 𒑕, and 𒑖 (contrast ordinary 𒈫 and 𒐈 otherwise used with 𒁹 numerals). As described in [Hue11, p. 585 n. (b) and (f)], the sign 𒄥 GUR, while it is used only with volumes in excess of one gur, is written after the whole expression, after the overt unit sign 𒋡 if present, and after the word for “grain” if present, as in

𒐞𒐎𒌋𒐉	𒑑	𒐋	𒋡	𒊺	𒄥[21]
3554 gur	3 ban$_2$	6	sila$_3$	of grain.	

Observe that while large numbers of gur follow[22] system $S_{\text{Ur III/OB}}$, the use of horizontal (AŠ) numerals for the gur disambiguates from the vertical bariga, as 𒌋𒁹𒄥 would be 10 gur 1 bariga, and 𒌋𒀸𒄥 would be 11 gur; again even with some overt units, most of the numerals that participate in a metrological system have an interpretation dependent on that system.

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

— 𒁹–𒐍 used in $S_{\text{Ur III/OB}}$ and the SPVS as well as with overt units;
— 𒌋–𒐐 used in $G_{\text{Ur III/OB}}$, of which 𒌋–𒐏 are also used in $S_{\text{Ur III/OB}}$ and the SPVS as well as with overt units;
— 𒁹–𒐎 used in $S_{\text{Ur III/OB}}$, and sometimes with overt units;
— 𒐞–𒐥 used in $S_{\text{Ur III/OB}}$;
— 𒊹–𒐦 used in $S_{\text{Ur III/OB}}$ and $G_{\text{Ur III/OB}}$;
— 𒐬–𒐭 used in $S_{\text{Ur III/OB}}$ and $G_{\text{Ur III/OB}}$;
— 𒀸–𒐋 used in $C_{\text{Ur III/OB}}$ as well as with overt units of the weight system;
— 𒑏, 𒑐, 𒑑, 𒑒, 𒑓 used in $C_{\text{Ur III/OB}}$;
— 𒁹, 𒑔, 𒑕, 𒑖 used in $C_{\text{Ur III/OB}}$—note the overlap with 𒁹–𒐋;
— 𒑘 and 𒑙 used in $G_{\text{Ur III/OB}}$.

Only in the SPVS did numerals exist truly independently of metrology; to quote [Rob08, p. 78]: “The SPVS temporarily changed the status of numbers from properties of real-world objects to independent entities that could be manipulated without regard to […] metrological system. […] Once the calculation was done, the result was expressed in the most appropriate metrological units and thus re-entered the natural world as a concrete quantity.”

[20] Used 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 $G_{\text{Ur III/OB}}$ based on a height of one cubit, see[Pow87, p. 488; Rob08, p. 294; Rob19].

[21] From [**P309594**].

[22] A larger unit, the guru$_7$ (*karûm*, grain heap), is sometimes used instead, with 𒀸𒂆𒋡𒐬𒌋=𒊹𒄥 (1 *karûm* = 3600 kurrū). See [Fri07, p. 415; Rob19].

3.3.4 The length system

In the Ur III and Old Babylonian periods, lengths are expressed using overt units counted with 𒁹- and 𒌋 numerals with their system $S_{\text{Ur III/OB}}$ values[23]. 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]:

𒁹𒆜	$\xleftarrow{60}$	𒁹𒍑	$\xleftarrow{10}$	𒁹𒃻	$\xleftarrow{12}$	𒁹𒌑	$\xleftarrow{30}$	𒁹𒋗𒋛.	$(L_{\text{Ur III/OB}})$
danna		UŠ[24]		nindan		kuš₃		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]:

𒁹𒆜	$\xleftarrow{30}$	𒁹𒍑	$\xleftarrow{6}$	𒁹𒂠	$\xleftarrow{10}$	𒁹𒃻	$\xleftarrow{2}$	𒁹𒄀	$\xleftarrow{6}$	𒁹𒌑	$\xleftarrow{30}$	𒁹𒋗𒋛.	$(\bar{L}_{\text{Ur III/OB}})$
				eše₂				gi					
				ašlum				*qânum*					
				rope				reed					
				60 m				3 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 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 $S_{\text{Ur III/OB}}$ [Fri07, p. 374; DE87, pp. 127, 165]:

$$◉ \xleftarrow{10} ● \xleftarrow{6} ◗• \xleftarrow{10} ◗ \xleftarrow{6} • \xleftarrow{10} ⊳. \qquad (S)$$

Likewise the area system used in the Early Dynastic IIIb period mirrors system $G_{\text{Ur III/OB}}$ [LAK, p. 72; NDE93, p. 63; Fri07, p. 378; Lec16]:

$$◉ \xleftarrow{10} ● \xleftarrow{6} ✹ \xleftarrow{10} • \xleftarrow{3} ⊳• \xleftarrow{6} ⊳, \qquad (G_{\text{ED IIIb}})$$

As noted in [L2/04-099, p. 4] (see section 3.2), the vertical 𒁹 from $S_{\text{Ur III/OB}}$ 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]:

$$● \xleftarrow{6} ◉ \xleftarrow{10} • \xleftarrow{3} ⊳• \xleftarrow{6} ⊳, \qquad (G)$$

Observe that, as noted in [DE87, p. 142], ◉ changes meaning from 10• in system G to 600• in system $G_{\text{ED IIIb}}$. 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 $G_{\text{ED IIIb}}$ is the "area 1" system).

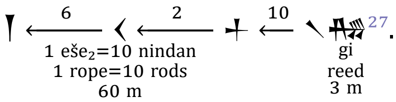
---

[23] Adjacent units are no more than a factor of 60 apart, so higher numerals such as 𒐕 or ◇ are not used.

[24] TODO

3.4.1 Field lengths in Ŋirsu

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



[27]. ($L_{\text{ED IIIb}}$)

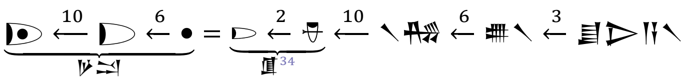
This is the system that was used to express the sides of the field in [P020054] discussed in section 3.2. In that tablet and others from the same period, such as the ones discussed in [Lec20], areas are expressed in system $G_{\text{ED IIIb}}$, with curviform numerals[28]; in the absence of overt units, such as when dealing with length that are integer multiples of a half-rope[29], 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 $G_{\text{ED IIIb}}$ and $L_{\text{ED IIIb}}$. The sign , 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 .

3.4.2 Dyke lengths in Ŋirsu

[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." [TODO(egg): also mention Krebernik 1998 p. 303 n. 686.] Powell does not elaborate on the specifics of this mixed use of numerals, but a cursory search in CDLI finds many occurrences[30], such as:

— [P221305, obv. 1, 4][31]
— [P020129, rev. 2, 1]
— [P221291, rev. 5, 1][32]
— [P221266, rev. 2, 1]

These expressions use an explicit sign (counted in multiples of ten) or . This notation—but not its use of curviform numerals—is remarked on in [Lec20, p. 290 n. 27], which cites several of the instances listed above. It seems to be typical of texts about dykes. These[33] can be summarized by the following factor diagram:



($L'_{\text{ED IIIb}}$)

[27] Note that the reeds are counted using *tenû* numerals, , , , etc.
[28] TODO(egg): Note the handful of late Urukagina tablets that start to have cuneiform areas.
[29] This is the case of the sides of the field in [P020054, obv. ii 2–3].
[30] A search for curviform numerals followed by some number of reeds counted in (*tenû*) cuneiform numerals currently finds 125 occurrences across 47 tablets.
[31] CDLI only has a copy, but a photo may be found in [Lec12, p. 82]. On that photo the is not visible. Lecompte notes that the copy is faithful; indeed another can be seen both on the copy and the photo on obv. 2, 2.
[32] From copy.
[33] TODO Cite also DP 568, the one with and even though it has no reeds.

3.4.3 Grain in Ŋirsu

TODO note the 𒋗𒀹 ⊂⊐▷. Also cite Friberg1978 p. 43 and reproduce its factor diagram.

3.4.4 Grain in Ebla

Lengths of Early Dynastic IIIb dykes from Ŋirsu are far from the only numeric expressions that mix curviform and cuneiform numerals.

The system of grain[35] capacities in Ebla uses the following units[36]:

$$\underset{\text{gu}_2\text{-bar}}{𒄘𒁇} \xleftarrow{2} \underset{\text{ba-ri}_2\text{-zu}}{𒁀𒌷𒍪} \xleftarrow{\frac{5}{2}} \underset{\text{ŋin}_4}{𒂅} \xleftarrow{4} \underset{\text{niŋ}_2\text{-sagšu}}{𒃻𒊕𒋗} \xleftarrow{6} \underset{\text{an-zam}_x}{𒀭𒍠}.$$

The 𒄘𒁇 and 𒁀𒌷𒍪 are generally counted using curviform numerals, and the smaller units using cuneiform 𒁹 numerals. Indeed, a search on [EbDA] for co-occurrences of either 𒀭𒍠 or 𒃻𒊕𒋗 with either of 𒄘𒁇 or 𒁀𒌷𒍪 finds the following expressions[37]:

1. [P240532, *verso* 4, 9] ◗⋘ 𒁀𒌷𒅗[39] 𒐊 𒃻𒊕𒋗
2. [P240548, *verso* 1, 1] ◗𒁀𒌷𒍪 𒐈 𒀭𒍠
3. [P240655, *recto* 7, 9] ◗◗⋘𒁇[40] 𒐊 𒃻𒊕𒋗
4. [P240579, *verso* 4, 3] ◗◗◗◗◗▽𒄘𒁇 𒈫 𒃻𒊕𒋗
5. [P240675, *verso* 2, 2] ◗⋘ 𒁀𒌷𒍪 𒐊 𒃻𒊕𒋗
6. [P240609, *verso* 3, 1] ◗𒁀𒌷𒍪 𒐈 𒀭𒍠
7. [P240533, *recto* 3, 3] ●●◗▽𒄘𒁇 𒐋 𒃻𒊕𒋗 𒈫𒈦 𒀭𒍠𒐉
8. [P240697, *recto* 1, 5] ◗▽[41] 𒈫𒂅 𒐈 𒀭𒍠
9. [P240653, *recto* 6, 2] ●●◗◗◗▽𒄘𒁇 𒁹𒃻𒊕𒋗 𒈫𒈦 𒀭𒍠
10. [P240654, *recto* 2, 6] ◗𒁀𒌷𒍪 𒐋𒋡[42] 𒐈 𒀭𒍠[43]

---

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

[35] Liquid capacities use a different system [Arc15, p. 229 n. 12]:

$$\underset{\text{la-ḫa}}{𒆷𒄩} \xleftarrow{30} \underset{\text{sila}_3}{𒋡} \xleftarrow{6} \underset{\text{an-zam}_x}{𒀭𒍠}.$$

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

$$\underbrace{◗⊂𒆷𒄩 \xleftarrow{\frac{5}{3}} ◗ \xleftarrow{6} ● \xleftarrow{10} ◗}_{𒆷𒄩} \xleftarrow{3} \underbrace{⟨ \xleftarrow{10} 𒁹}_{𒋡} \xleftarrow{6} 𒁹𒀭𒍠,$$

but we have not investigated this thoroughly.

[36] TODO mention the other one citing Chambon and the footnote in Archi

[37] We cite here only one attestation per tablet; most tablets contain several expressions mixing curviform 𒁀𒌷𒍪 and larger with cuneiform 𒂅 and smaller. In all cases the transcriptions given here are based on the EbDA transliterations, but the shape and orientation of the numerals was checked[38] on a photograph (from EbDA unless noted otherwise).

[38] As 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.

[39] ba-ri$_2$-zu$_2$, a variant spelling.

[40] Short for 𒄘𒁇.

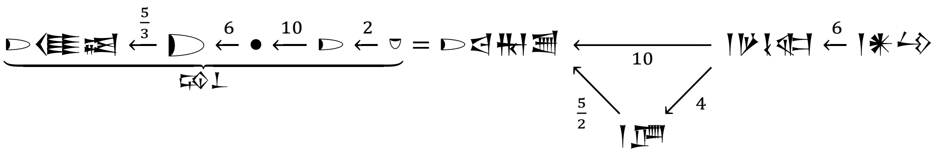
[41] Note the omitted 𒄘𒁇.

[42] Instead of the expected 𒃻𒊕𒋗.

[43] 𒐈𒀭𒍠 not legible on the EbDA photo.

11. [P240531, *recto* 1, 8] [cuneiform]
12. [P241708, *recto* 1, 1][44] [cuneiform]
13. [P241904, *recto* 1, 1][45] [cuneiform][46][cuneiform]

Note that higher numbers of [cuneiform] are expressed in hundreds (*mi-at* [cuneiform]) and then thousands (*li-im* [cuneiform]), as is typical in Ebla [Arc15, p. 33], *e.g.*, in [P240532, *verso* 2, 3], [cuneiform] $(100+60+30+5 = 195$ [cuneiform] of grain). These expressions correspond to the following factor diagram:



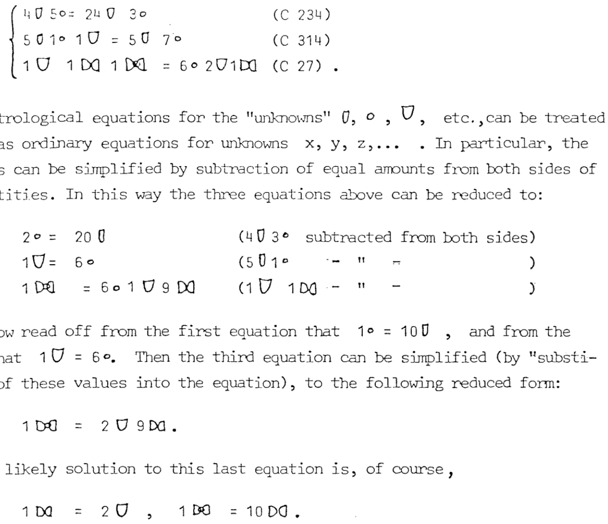
(C_{Ebla})

3.4.5 Use in modern publications

Because of their prevalence in the Uruk and Early Dynastic periods, the proposed numerals are widely used in modern publications discussing metrology in those periods, as illustrated in Figures 3–15.

Since they contrast with the cuneiform numerals, they likewise appear contrastively in such publications. A remarkable example of that is found in Figure 15. The partial[47] transliteration "4[cuneiform] *ʾa₃-da-um* 4[cuneiform] *aktum* 4[cuneiform] $ib_2^{tu_9}$×3[cuneiform] sa_6 $gunu_3$" is used to illustrate a discussion of the interpretation of the contrast between [cuneiform] and [cuneiform] numerals. More conventional transliterations[48] might omit the numeral shapes entirely, *e.g.*, 4 *ʾa₃-da-um* 4 *aktum* 4 $ib_2^{tu_9}$×3 sa_6 $gunu_3$, which would obviously be inadequate in this context. There are transliteration conventions that are more explicit about numeral shape, *e.g.*, 4(aš^c) *ʾa₃-da-um* 4(diš^c) *aktum* 4(aš^c) $ib_2^{tu_9}$×3(diš) sa_6 $gunu_3$, but the result would be less readable. See Section 3.7.2 for a discussion of transliteration conventions for numerals.

[44] From CDLI photo.
[45] From photo in [Arc89, p. 6].
[46] Laid out as [cuneiform]; on stacking patterns see Section 6.3.
[47] The untransliterated text would be [cuneiform]; note the atomically encoded $ib_2 \times 3$[cuneiform] = [cuneiform] × [cuneiform] = [cuneiform].
[48] TODO cite the EbDA one.
[49] TODO(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.



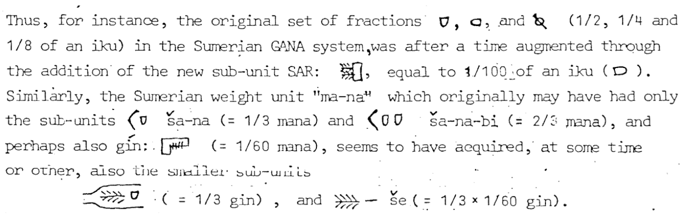
(C 234)
(C 314)
(C 27) .

These metrological equations for the "unknowns" etc., can be treated exactly as ordinary equations for unknowns x, y, z,... . In particular, the equations can be simplified by subtraction of equal amounts from both sides of the identities. In this way the three equations above can be reduced to:

We can now read off from the first equation that, and from the second that. Then the third equation can be simplified (by "substitution" of these values into the equation), to the following reduced form:

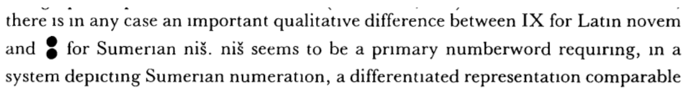
The most likely solution to this last equation is, of course,

Figure 1: TODO [Fri78, p. 15]



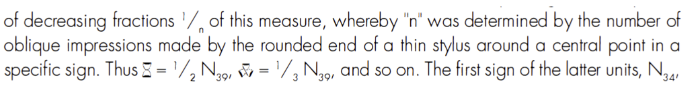
Thus, for instance, the original set of fractions (1/2, 1/4 and 1/8 of an iku) in the Sumerian GANA system, was after a time augmented through the addition of the new sub-unit SAR: equal to 1/100 of an iku. Similarly, the Sumerian weight unit "ma-na" which originally may have had only the sub-units ša-na (= 1/3 mana) and ša-na-bi (= 2/3 mana), and perhaps also gín: (= 1/60 mana), seems to have acquired, at some time or other, also the smaller sub-units (= 1/3 gin), and še (= 1/3 × 1/60 gin).

Figure 2: TODO [Fri78, p. 49]



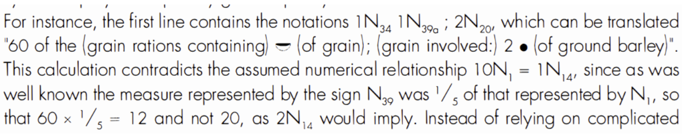
there is in any case an important qualitative difference between IX for Latin novem and for Sumerian niš. niš seems to be a primary numberword requiring, in a system depicting Sumerian numeration, a differentiated representation comparable

Figure 3: TODO [**Englund1988**]



of decreasing fractions $^1/_n$ of this measure, whereby "n" was determined by the number of oblique impressions made by the rounded end of a thin stylus around a central point in a specific sign. Thus $= {}^1/_2\ N_{39}$, $= {}^1/_3\ N_{39}$, and so on. The first sign of the latter units, N_{34},

Figure 4: TODO [Eng98, p. 113]

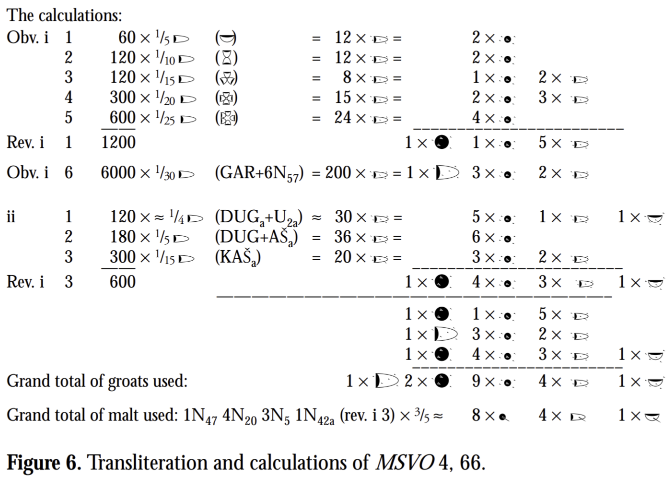


For instance, the first line contains the notations $1N_{34}\ 1N_{39a}$; $2N_{20}$, which can be translated "60 of the (grain rations containing) (of grain); (grain involved:) 2 • (of ground barley)". This calculation contradicts the assumed numerical relationship $10N_1 = 1N_{14}$, since as was well known the measure represented by the sign N_{39} was $^1/_5$ of that represented by N_1, so that $60 \times {}^1/_5 = 12$ and not 20, as $2N_{14}$ would imply. Instead of relying on complicated

Figure 5: TODO [Eng98, p. 116]

Die halbkreisförmigen Griffeleindrücke gehen manchmal in mehr oder weniger eckige Formen über ( )[685]. Es gibt aber auch Einer in Form von regelrechten – meist mehr oder weniger schräggestellten – Keilen ( ), die öfters neben halbrunden Einern vorkommen und mit diesen kontrastieren[686]. Selten treten mit gebildete Zahlen auf[687] (sie entsprechen den bariga-Zahlen im Hohlmaßsystem, s.u. 7.4).

Figure 6: TODO [Kre98, p. 303]



The calculations:

Obv. i	1	60 × 1/5		=	12 × =		2 ×		
	2	120 × 1/10		=	12 × =		2 ×		
	3	120 × 1/15		=	8 × =		1 ×	2 ×	
	4	300 × 1/20		=	15 × =		2 ×	3 ×	
	5	600 × 1/25		=	24 × =		4 ×		
Rev. i	1	1200				1 ×	1 ×	5 ×	
Obv. i	6	6000 × 1/30	(GAR+6N57)	=	200 × =	1 ×	3 ×	2 ×	
ii	1	120 × ≈ 1/4	(DUGa+U2a)	≈	30 × =		5 ×	1 ×	1 ×
	2	180 × 1/5	(DUG+AŠa)	=	36 × =		6 ×		
	3	300 × 1/15	(KAŠa)	=	20 × =		3 ×	2 ×	
Rev. i	3	600				1 ×	4 ×	3 ×	1 ×
						1 ×	1 ×	5 ×	
						1 ×	3 ×	2 ×	
						1 ×	4 ×	3 ×	1 ×
Grand total of groats used:					1 ×	2 ×	9 ×	4 ×	1 ×
Grand total of malt used: 1N47 4N20 3N5 1N42a (rev. i 3) × 3/5 ≈							8 ×	4 ×	1 ×

Figure 6. Transliteration and calculations of *MSVO* 4, 66.

Figure 7: TODO [**Englund2001**]

strong similarities between "area" 1 and "area" 3 systems, the sign with two concentric discs (, notated N_{50}[27]) remains problematic. It never appears in any numerical combination with the sign with a single disc (,

Figure 8: TODO [Cha03, p. 6]

formed by only two signs 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, 60^2, 60^3, 1/60, $1/60^2$, etc. are written in the same way, with the vertical wedge . The third feature concerns the exact function of

Figure 9: TODO [Cha12, p. 58]

one step. The scribes of the Early Dynastic Period (*c.* 2600 BC), for instance, represented the number 648, 000 with: but never with the repetition .

Figure 10: TODO [Cha12, p. 59][49]

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:

še bar *ba-rí-zu*

'3 *gubar* (capacity units) and 1 *parīsu*'.

Figure 11: TODO [Cha12, p. 61]

This is particularly true of the signs , , and , 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 , , and refer to 1, 2, 3,

Figure 12: TODO [Cha12, p. 64]

shape. The principle of notation is additive: each sign is noted as many times as necessary (e.g., transliterated as 2(šar$_2$) 1(geš'u) 3(u), means 2 × 3600 + 1 × 600 + 3 × 10). The system is based on an alternation of factors ten and

Figure 13: 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

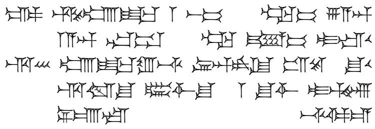
Fig. 1. Combinations of numerals attested in Š. 742.

Figure 14: 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 *ʾa$_3$-da-um*^{tu9}-2 4 aktum 4 ib$_2$^{tu9} ×3 sa$_6$ gunu$_3$" (Fig. 2).

Figure 15: 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’[50]. Do you know its reading[51]?

Examenstext A

Many of the cuneiform numerals are used with a logographic or phonetic value. For example, the sign 𒀸 has, *inter alia*, the values aš, rum, and dili. While the horizontal numerals are most frequently written ⊏ in the Early Dynastic period[52], such non-numeric usage is almost[53] always written 𒀸, for instance:

- in personal names in administrative texts, such as the following, which all contain ⊏ numerals:
  - 𒀸𒀵 in [**P010424**; **P010458**; **P010459**] from ED IIIa أبو صلابيخ,
  - 𒀸𒋛 in [**P010960**] from ED IIIa Šuruppag,
  - 𒈬𒌋𒀸𒂗 in [**P251641**] from ED IIIb Adab,
  - 𒄑𒈠𒀸𒂗𒁹 in [**P252866**] from ED IIIb Adab,
  - 𒂊𒋛𒂊𒀸𒋛 in [**P298637**] from ED IIIb Umma;
- in the Sumerian word 𒌑𒀸 u$_2$-rum, “property” in ED IIIb Ŋirsu administrative texts which contain ⊏ numerals, such as [**P020006**; **P020008**; **P020018**; **P020024**; **P020030**];
- in lexical texts:
  - in the divine name 𒀭𒉌𒃲𒀸𒆠 in the lexical texts [**P010570**; **P010572**], where the entries are prefixed with ⊏.
  - in the word 𒀸 dili, “small fish” in [**P010578**], witness to Early Dynastic Fish,
  - in the same word with a determinative, 𒀸𒄩 dili^{ku_6}, in [**P010586**], witness to Early Dynastic Food, which starts with ⊏ numerals.

This is a clear contrast between 𒀸 and ⊏ in this period, and genuine ambiguity can arise if it is lost; for instance, the personal name 𒀸𒀵 occurs on its own line in the aforementioned administrative texts; a line ⊏𒀵 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

[50]The reader will recall that ŋeš$_2$ is written 𒐕, with a larger wedge than 𒁹; however, these signs have merged by the time Examenstext A is composed.

[51]Besides ŋeš$_2$, a look at [OSL] shows that the values diš, ge$_3$, makkaš, saŋtak$_4$, and tal$_4$ are attested both in [ePSD2] and in lexical lists. The sign is also used for the Akkadian word *ana* in the Neo-Assyrian period.

[52]A 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)"`.

[53]Exceptions are discussed in section 3.7.1.

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 𒀸 that have both numeric and non-numeric values in both the main list and the section on numbers. For instance, [KWU, pp. 123 sqq.] lists all of 𒀸–𒐂 together with ⊃–𒐎, while 𒀸, 𒐀, and 𒐁, and only those, appear at the beginning of the sign list, since they have non-numeric values[54]. [PTACE, p. 58] has the numeric signs ⊃, 𒀹, 𒁹, whereas non-numeric 𒀸 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, [LAK] writes *s. die Zahlz.* throughout the main list; LAK 1 𒀸 thus reappears at LAK 829 together with ⊃, 𒀹, and 𒁹. One should note [MZL], 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 𒐛–𒐝 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 changes[55] to the text between Early Dynastic and later witnesses that prevent a diachronic encoding of such composites. For numerals, the switch from ⊃ to 𒁹 numerals prevents diachronic encoding even if ⊃ were unified with 𒀸. 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 𒁹 numerals, whereas they are written with ⊃ numerals in the Early Dynastic witnesses; since 𒁹 and 𒀸 are distinct[56] characters, the ⊃-𒀸 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 periods[57], 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

[54]Non-numeric values of 𒀸 were discussed in section 3.5; 𒐀 has the values man₃ and min₅, and is used for the word didli, "several, various"; 𒐁 has the value eš₆.

[55]TODO comment on the ED witnesses to the instructions of Šuruppag

[56]Besides 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 𒀸 which are not shared with 𒁹 have already been mentioned, but perhaps most striking is the fact that, in the Neo-Assyrian period, 𒀸 is used for the preposition *ina*, "in", and 𒁹 for the preposition *ana*, "to".

[57]TODO cite a few things here.

in [UTR56, §2.4]. The lack of dedicated Ur III fonts may be explainable by the chart-like fonts[58] 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 ŠAR$_2$

The character U+122B9 ● CUNEIFORM SIGN SHAR2 has a circular reference glyph.

In most texts from the Early Dynastic IIIb and Old Akkadian period[59], a contrast between non-numeric šar$_2$ written ◇ and numeric 1(šar$_2$^{c}) written ● can be observed, similar to the contrast between 𒀸 and ⊃ previously discussed in section 3.5. However, in lexical lists from Šuruppag and Ebla[60], as well as in the *Stèle des vautours*, non-numeric šar$_2$ is curviform:

— 𒀭 [illegible] ● and 𒀭 [illegible] ● [illegible] in [**P010566**];
— ● [illegible] and 𒀭 ● [illegible] in [**P010576**];
— ● [illegible] in [**P240986**][61];
— ● [illegible] in [P222399, obv. 17, 9, 18, 11, 22, 12][62].

It *would* be disruptive to the diachronic representation of text if non-numeric šar$_2$ 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 šar$_2$ 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 𒀸 (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

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

[59]For example, in personal names:

— [illegible] ◇ [illegible] in [**P020019**] from ED IIIb Ŋirsu;
— [illegible] ◇ [illegible] in [**P020182**], also from ED IIIb Ŋirsu;
— [illegible] 𒀭 ◇ in [**P222186**] from ED IIIb Umma;
— [illegible] 𒀭 [illegible] ◇ in [**P235312**] from Old Akkadian Umma.

[60]TODO Mention other ways in which these are archaizing

[61]From copy in [ELLes, No. 397].

[62]Note however 𒀭 [illegible] ◇ [illegible] on [P222399, obv. 6, 17]. Curviform non-numeric šar$_2$ is clearly archaizing in ED IIIb Ŋirsu; one might suppose that the scribe slipped into their modern ways here. TODO add a photo.

ePSD2 one (uru$_8$^{rum}) of this word disagree on that aspect. Since 𒀸 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 BIL$_2$ can be described as NE×KASKAL or NE×PAP[63], they are typically all transliterated bil$_2$, and therefore are all represented by the character U+1224B 𒉋 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, the sequence 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 sequence from [P240533, *recto* 3, 3] is transliterated “20-1-1/2 *gu*$_2$*-bar* 7 nig$_2$-sagšu 2-1/2 an-zam$_x$[64] *za*”, reading both of its 1/2 signs as 1/2, but not distinguishing them.

In particular, these transliterations do not differentiate between 𒀸 and 𒁹 numerals, nor between the two curviform and cuneiform 10 numerals. For instance, the aforementioned sequence from [P242293, *recto* 4, 1] is transliterated “4 *ʾa*$_3$*-da-um*^{tug_2}-II 4 aktum^{tug_2} 4 ib$_2$-III gun$_3$ sa$_6$ ^{tug_2}” in EbDA, with no distinction between the two forms of 4. Since 𒀸 and 𒁹 numerals are separately encoded, the numeric expressions in such transliterations cannot be transformed into Unicode cuneiform without additional context, regardless of curviform–cuneiform unification.

In metrological systems such as systems $G_{\text{Ur III/OB}}$ and $C_{\text{Ur III/OB}}$ 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, the sequence from [**P386847**] is transliterated “1(eše$_3$) 5½ iku[65] 7 sar” in [Feu04, vol. 2, p. 176], and the sequence from [**P307255**] is transliterated “1(n[66]) 2(b) 7 ½ sila$_3$” 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 the above examples are “1(gesz2) 3(u) 5(asz) gur”, “1(esze3) 5(iku) 1/2(iku) GAN2 7(disz) sar”, and “3(barig) 2(ban2) 7(disz) 1/2(disz) sila3”. CDLI and ePSD2 both distinguish curviform from cuneiform numerals in transliteration: the length 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(geš$_2$^{c}) 3(u^c) ^{ninda}ninda$_x$(DU) 1/2(aš^c) 4(diš^t) gi” in ePSD2. Another example is [**Molina2014**], which uses 1*a* for 𒀸, 1*d* for 𒁹, 1*ac* and 1*dc* or ½*dc* 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.

[63]As on [P249253].

[64]As of this writing, EbDA actually has an-zam$_\chi$, with U+1D6A GREEK SUBSCRIPT SMALL LETTER CHI.

[65]TODO say something about this reading

[66]TODO comment on nigida.

While there exist transliterations that distinguish from but not from , such as the ones used in [**DCCMT**], the trend, especially in more recent works in third millenium 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 , with a c indicating that the GUR numerals are curviform, and the parenthetical GUR indicating that these are rather than numerals.

3.8 Conclusions

Co-occurences of curviform and cuneiform numerals are not anecdotal in the Early Dynastic period, nor are they the result of scribal idosyncrasy. Instead, they represent systematic 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 would require any treatment of Early Dynastic administrative texts to use multiple cuneiform fonts, often within single numeric expressions. Further, if that contrast is lost in plain-text interchange, the text can be misinterpreted: is a length of three ropes, but is an area of three bur$_3$; could be read as one and one , where would be one and a half ; is a personal name, but 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 millenium corpora. The oddity of ● requires some explanation, but does not 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

A complete rationale for disunification between the non-numeric signs used in the fourth millenium and the already-encoded cuneiform signs will be given in the forthcoming proto-cuneiform encoding proposal. The core issue with extending the cuneiform script further back in time is that, since 1987, fourth millenium studies have used a different model of character identity and associated transliteration conventions, with names being given to structurally different glyphs, and no attempt being made at assigning phonetic values to them.

This is not a mere classification of glyph variants, as contrastive meanings of these systematic variants can often be reconstructed, with, *e.g.*, signs KAŠa, KAŠb, and KAŠc, depicting filled jars with a spout (a), a handle (c), or neither (b), being understood as referring to containers of different substances, see [**Englund2001**]. However, not all identified systematic variants are understood, and the general approach to character identity is closer to that used for undeciphered or partially deciphered script.

As part of the development of these conventions, a classification of fourth millenium numeric signs was developed; see [DE87]. This classification assigns to each unit numerals an identifier formed by the letter *N* with a numeric subscript (sometimes with an additional alphabetic subscript): N_1 is 𒀸, N_{14} is ●, N_{34} is 𒀸, etc. Transliterations of numeric expression then use those to identify the type of number used, thus $5N_1$ is 𒐊, and $5N_{14}$ is ●●●●●.

In contrast with the use of parenthetical unit names, this approach does not require interpreting the quantity being counted. This is valuable in contexts where the exact meaning of the choice of numeral type is not fully understood, as the transliterations can otherwise force a dubious interpretation. For instance, the CDLI transliteration of 𒀸𒀸𒀸 𒁀 or 𒐈 𒁀 in [**P283802**] currently uses (barig@c) for the vertical numerals, suggesting a capacity measure; but [Gor23] interprets these instead as counting linen textiles. As a result, the fourth millenium conventions for numeral transliteration are used in Early Dynastic texts, especially those from the ED I–II period, even though the Sumerian text uses classical assyriological transliteration conventions.

While the non-numeric signs are treated as undeciphered, the metrological systems used in the fourth millenium are well understood, as can be seen in [DE87, p. 165]. As a result, contrary to the non-numeric proto-cuneiform conventions, these numeric transliteration conventions are compatible with the classical ones described in section 3.7.2; they are indeed used interchangeably, as in [**P011104**] which uses the notation u@f in [ePSD2], but N14@f in CDLI. Indeed, the numerals are used similarly in Early Dynastic metrological systems, and are visually identical.

A disunification of numerals between the third and fourth millenium would therefore induce confusion as to which numerals should be used in third millenium studies, and would needlessly duplicate the encoding of at least seventy characters; by splitting the attestations, these separate encoding proposals would run into additional difficulties to supply evidence for encoding.

Note that the structural variants designated by letters in fourth millenium notation have systematically been encoded, as they have occasionally be found to carry distinct numeric meaning. For instance, 𒋰 N_{30c} is listed as a variant of 𒋰 N_{30a} in [DE87, p. 166], where the numeric value of either in relation to 𒌋 N_{29a} is still unknown, but their values are found in [**Englund2004**] to be 𒋰 $= \frac{1}{10}$𒌋, whereas 𒋰 $= \frac{1}{6}$𒌋.

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(ban$_2$@c) and 5(ban$_2$@c).]

6 Characters not included in this proposal

6.1 The fractions $\frac{1}{3}$ and $\frac{2}{3}$

TODO 𒑚, 𒑛. Note the occasional omission of 𒁹, see citations in email to steve plus https://cdli.mpiwg-berlin.mpg.de/artifacts/274845/reader/51537.

6.2 Missing numerals

TODO N13 not attested in CDLI TODO (N_{17} not usefully numeric, $12N_{14}$ not encodable, etc.). Cite [DE87, p. 147] 7 and 8(diš *tenû*) encodable, but not today; want to go

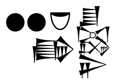


Figure 16: The layout of case [**P011099**]; the numeral ⁚ is rotated to fit the rounded corner of the tablet.

into the Cuneiform Numbers and Punctuation block for sanity.

6.3 Stacking patterns

The already-encoded numerals in the Cuneiform Numbers and Punctuation block distinguish some *stacking patterns*; for instance 9𒁹 is encoded both as U+12446 𒑆 and as U+1240E 𒐎. This is in part due to contrastive usage of stacking patterns. For instance, besides 𒑘 and 𒑙 which are characteristic of bariga measures, four bariga is written 𒐼 even where 4𒁹 is written 𒐉, as in [**P255010**; **P292843**]. Another contrast is that between the stacking patterns used in scratch calculations in the SPVS, often 𒁹 𒈫 𒐈 𒐉 𒐊 𒐋 𒑂 𒑄 𒑆 𒌋 𒎙 𒌍 𒑩 𒑪, and results in metrological systems, typically 𒁹 𒈫 𒐈 𒐼 𒐊 𒐋 𒐌 𒐍 𒐎 𒌋 𒎙 𒌍 𒐏 𒐐, occasionally co-occurring as in [**P142827**; **P142357**]. This separate encoding is also for compatibility with distinctions made in reference works and in some non-numeric transliterations; for instance, 𒐉 is [MZL, No. 860] and has the value limmu, and 𒐼 is [MZL, No. 852] and has the value limmu_5. Numeric[67] transliterations occasionally distinguish the stacking patterns 𒐉 𒑂 𒑄 𒑆 𒑩 𒑪, as in the CDLI transliterations of the aforementioned tablets, although this is rare; often 4(diš) is 𒐼 in Ur III, but 𒐉 in the Neo-Assyrian period.

However, the stacking patterns from earlier periods are not separately encoded; for instance, in ED IIIb Ŋirsu, 𒎙 2(u) often has one 𒌋 atop another. These older stacking patterns do not appear to be contrastive, are not marked in transliteration, and are not listed separately in sign lists nor assigned any different values. There is therefore no evidence of a need to encode them; instead, they should be considered style variants, and an ED IIIb Ŋirsu font should have an appropriate glyph for U+12399 𒎙 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; these distinctions would be incompatible with the state of the art in numeric transliterations, and are not needed to represent reference works. 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 numerals may be laid out across the case in whichever way fits the available space; this is illustrated in Figure 16. Note also that the numerals need to be considerably enlarged in order to reproduce the layout of the tablets, so that ⁚ often spans two lines of cuneiform signs, as shown in Figure 17. 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 Figure 3: reproducing the layout of tablets is not within the scope of plain text.

[67] The Sumerian word limmu means "four", so limmu and limmu_5 are still numbers. The distinction here is between usage in transliterations of phrases such as 𒈗 𒀭𒌒𒁕 𒐼𒁀𒆤 lugal an-ub-da limmu_5-ba-ke_4 (king of the four quarters) or of names, and of numeric expressions such as 𒐼𒋡 4(diš) sila_3.

Figure 17: 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".

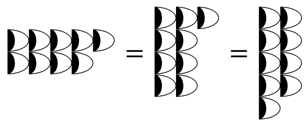


Figure 18: Three stacking patterns for U+12573 CUNEIFORM NUMERIC SIGN NINE N34. The one on the left is the reference glyph, used in Uruk III [**P003499**; **P004430**], and widely afterwards, *e.g.*, ED IIIa Šuruppag [**P010678**], ED IIIb Ŋirsu [**P020057**], Old Akkadian Umma [**P212464**]. The ones in the middle and right are used in two Uruk IV tablets [**P001243**; **P004500**]. All three Uruk examples are transliterated 9(N34) in CDLI.

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 18. 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 18 uses OpenType stylistic alternates, and Figure 16 rotates the character , in both cases preserving the plain text backing.

6.4 Other glyph variants not reflected in transliteration

TODO Comment on the nameless variant glyphs from L2/23-190 and note that they are illustrating an even wider glyphic range as shown in [**Englund2001**].

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/.

Major reference works and online projects

[EbDA] 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/.

[eBL] E. Jiménez, Z. Földi, A. Hätinen, 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/.

[ELLes] P. Mander. "Lista dei segni dei testi lessicali di Ebla". In: *Testi lessicali monolingui della biblioteca L. 2769*. Ed. by G. Pettinato. Materiali epigrafici di Ebla 3. Napoli: Istituto universitario orientale, 1981, pp. 285–382.

[ePSD2] S. Tinney, P. Jones and N. Veldhuis, eds. *The electronic Pennsylvania Sumerian Dictionary*. 2nd ed. 2017–.
eprint: http://oracc.org/epsd2.

[KWU] N. Schneider. *Die Keilschriftzeichen der Wirtschaftsurkunden von Ur III*. Editrice Pontificio Istituto Biblico, 1935.

[LAK] A. Deimel. *Liste der archaischen Keilschriftzeichen von Fara*. Wissenschaftliche Veröffentlichungen der Deutschen Orient-Gesellschaft 40. J. C. Hinrichs'sche Buchhandlung, 1922.

[MZL] R. Borger. *Mesopotamisches Zeichenlexikon*. Alter Orient und Altes Testament 305. Ugarit-Verlag, 2010.

[OSL] N. Veldhuis, S. Tinney et al., eds. *Oracc Sign List*. 2014–.
eprint: http://oracc.org/osl/.

[PTACE] A. Catagnoti. *La paleografia dei testi dell'amministrazione e della cancelleria di Ebla*. Quaderni di Semitistica 9. Università di Firenze, 2013.
ISBN: 8890134054.

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.

[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.

[Cap02] R. Caplice. *Introduction to Akkadian*. 4th ed. Editrice Pontificio Istituto Biblico, 2002.
ISBN: 88-7653-566-7.

[Cha03] 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.

[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.

[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.

[Gor23] F. Gori. "On Lapis Lazuli and Linen in Šuruppag 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.

[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. Krebernik. "Die Texte aus Fāra und Tell Abū Ṣalābīḫ". 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 Ĝirsu et d'une notation numérique inhabituelle". In: *Altorientalische Forschungen* 39.1 (Dec. 2012), pp. 81–86. DOI: 10.1524/aofo.2012.0006.

[Lec16] C. Lecompte. "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.

[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 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.

[Rob22] E. Robson. "Overview of Metrological Systems". In: *The Digital Corpus of Cuneiform Mathematical Texts*. 2022. ORACC: 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.

[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/RaiAbstractBookAjoitettuJaPäivätty_1.pdf.