

TACE16

Tamil: U+E100 – U+E3FF

Tamil All Character Encoding (TACE16) is a 16-bit Unicode-based character encoding scheme for Tamil language. Borne out of the discontent with the present Unicode for Tamil language due to unnatural assignments of Code Points and issues in handling Natural Language Processing tasks.

All characters of this encoding scheme are located in the *PRIVATE USE AREA* of the **Basic Multilingual Plane** (BMP) of Unicode's Universal Character Set. The Collation is sequential in accordance with the code point values and TACE16 encoding is unambiguous.

TACE16 follows the basic Tamil grammar rule: *மெய் + உயிர் = உயிர்மெய்* { Consonant + Vowel = Vowel-Consonant }

In TACE16, using the Vowels and Consonants, Vowel-Consonants can easily be derived by following one of the methods listed below.

Simple Arithmetic :

$$\begin{aligned} \text{க்} + \text{இ} &= \text{கி} \\ E210 + E203 - E200 &= E213 \end{aligned}$$

E200 is a constant.

Bitwise operation:

$$\begin{aligned} \text{க்} \mid (\text{இ} \& 000F) &= \text{கி} \\ E210 \mid (E203 \& 000F) &= E213 \end{aligned}$$

000F is the Bit-mask.

Bitwise operation(contd.):

/ To get Vowel from Vowel Consonant, Bit-mask F20F */*

$$\begin{aligned} \text{கி} \& F20F &= \text{இ} \\ E213 \& F20F &= E203 \end{aligned}$$

/ To get Consonant from a Vowel Consonant, Bit-mask FFF0 */*

$$\begin{aligned} \text{கி} \& FFF0 &= \text{க்} \\ E213 \& FFF0 &= E210 \end{aligned}$$

Bitwise operation(contd.):

/ Get Code Point of Tamil Number from English number */*

$$\begin{aligned} 5 \mid 0 &= 5 \\ 5 \mid E180 &= E185 \end{aligned}$$

/ Get English Number from Tamil Number */*

$$\begin{aligned} \text{௫} \& 000F &= 5 \\ E185 \& 000F &= 5 \end{aligned}$$

TABLE I: TAMIL SYMBOLS, NUMBERS AND FRACTIONS

௨ E100	மீ E101	ஹ் E102	யு E103	ஶ E104	ஷ் E105	ரீ E106	நீ E107	ஃ E108	● E109	★ E10A	☞ E10B	☹ E10C			
௦ E180	௧ E181	௨ E182	௩ E183	௪ E184	௫ E185	௬ E186	௭ E187	௮ E188	௯ E189	௩௦ E18A	௩௧ E18B	௩௨ E18C			
$\frac{1}{8}$ E1A0	$\frac{1}{4}$ E1A1	$\frac{1}{2}$ E1A2	$\frac{3}{4}$ E1A3	$\frac{1}{32}$ E1A4	$\frac{1}{16}$ E1A5	$\frac{3}{16}$ E1A6	$\frac{1}{40}$ E1A7	$\frac{1}{20}$ E1A8	$\frac{1}{10}$ E1A9	$\frac{3}{20}$ E1AA	$\frac{1}{250}$ E1AB	$\frac{1}{320}$ E1AC	$\frac{1}{160}$ E1AD	$\frac{1}{80}$ E1AE	$\frac{3}{80}$ E1AF

LEGEND:

ஃ	●	★	☞
பௌர்ணமி	அமாவாசை	கார்த்திகை	ராஜ

$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{32}$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{40}$
அரைக்கால்	கால்	அரை	முக்கால்	அரைவீசம்	வீசம்	மூவீசம்	அரைமா
$\frac{1}{20}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{250}$	$\frac{1}{320}$	$\frac{1}{160}$	$\frac{1}{80}$	$\frac{3}{80}$
ஒருமா	இரண்டுமா	மும்மா	நாலுமா	முந்திரி	அரைக்காணி	காணி	முக்காணி

TABLE 2: TAMIL VOWEL SIGNS AND VOWELS

ஂ E1F0	
	அ E201
ா E1F2	ஆ E202
ி E1F3	இ E203
ீ E1F4	ஈ E204
ஃ E1F5	உ E205
ஔ E1F6	ஊ E206
ெ E1F7	எ E207
ே E1F8	ஏ E208
ை E1F9	ஐ E209
ொ E1FA	ஓ E20A
ோ E1FB	ஔ E20B
ௌ E1FC	ஔ E20C
	ஃ E20D

TABLE 3: TAMIL CONSONANTS

க்	ங்	ச்	ஞ்	ட்	ண்	த்	ந்	ப்	ம்	ய்	ர்	ல்	வ்	ழ்	ள்	ற்	ன்
E210	E220	E230	E240	E250	E260	E270	E280	E290	E2A0	E2B0	E2C0	E2D0	E2E0	E2F0	E300	E310	E320
க	ங	ச	ஞ	ட	ண்	த	ந	ப	ம	ய	ர	ல	வ	ழ	ள	ற	ன
E211	E221	E231	E241	E251	E261	E271	E281	E291	E2A1	E2B1	E2C1	E2D1	E2E1	E2F1	E301	E311	E321
கா	ஙா	சா	ஞா	டா	ணா	தா	நா	பா	மா	யா	ரா	லா	வா	ழா	ளா	றா	னா
E212	E222	E232	E242	E252	E262	E272	E282	E292	E2A2	E2B2	E2C2	E2D2	E2E2	E2F2	E302	E312	E322
கி	ஙி	சி	ஞி	டி	ணி	தி	நி	பி	மி	யி	ரி	லி	வி	ழி	ளி	றி	னி
E213	E223	E233	E243	E253	E263	E273	E283	E293	E2A3	E2B3	E2C3	E2D3	E2E3	E2F3	E303	E313	E323
கீ	ஙீ	சீ	ஞீ	டீ	ணீ	தீ	நீ	பீ	மீ	யீ	ரீ	லீ	வீ	ழீ	ளீ	றீ	னீ
E214	E224	E234	E244	E254	E264	E274	E284	E294	E2A4	E2B4	E2C4	E2D4	E2E4	E2F4	E304	E314	E324
கு	ஙு	சு	ஞு	டு	ணு	து	நு	பு	மு	யு	ரு	லு	வு	ழு	ளு	று	னு
E215	E225	E235	E245	E255	E265	E275	E285	E295	E2A5	E2B5	E2C5	E2D5	E2E5	E2F5	E305	E315	E325
கூ	ஙூ	சூ	ஞூ	டூ	ணூ	தூ	நூ	பூ	மூ	யூ	ரூ	லூ	வூ	ழூ	ளூ	றூ	னூ
E216	E226	E236	E246	E256	E266	E276	E286	E296	E2A6	E2B6	E2C6	E2D6	E2E6	E2F6	E306	E316	E326
கெ	ஙெ	செ	ஞெ	டெ	ணெ	தெ	நெ	பெ	மெ	யெ	ரெ	லெ	வெ	ழெ	ளெ	றெ	னெ
E217	E227	E237	E247	E257	E267	E277	E287	E297	E2A7	E2B7	E2C7	E2D7	E2E7	E2F7	E307	E317	E327
கே	ஙே	சே	ஞே	டே	ணே	தே	நே	பே	மே	யே	ரே	லே	வே	ழே	ளே	றே	னே
E218	E228	E238	E248	E258	E268	E278	E288	E298	E2A8	E2B8	E2C8	E2D8	E2E8	E2F8	E308	E318	E328
கை	ஙை	சை	ஞை	டை	ணை	தை	நை	பை	மை	யை	ரை	லை	வை	ழை	ளை	றை	னை
E219	E229	E239	E249	E259	E269	E279	E289	E299	E2A9	E2B9	E2C9	E2D9	E2E9	E2F9	E309	E319	E329
கொ	ஙொ	சொ	ஞொ	டொ	ணொ	தொ	நொ	பொ	மொ	யொ	ரொ	லொ	வொ	ழொ	ளொ	றொ	னொ
E21A	E22A	E23A	E24A	E25A	E26A	E27A	E28A	E29A	E2AA	E2BA	E2CA	E2DA	E2EA	E2FA	E30A	E31A	E32A
கோ	ஙோ	சோ	ஞோ	டோ	ணோ	தோ	நோ	போ	மோ	யோ	ரோ	லோ	வோ	ழோ	ளோ	றோ	னோ
E21B	E22B	E23B	E24B	E25B	E26B	E27B	E28B	E29B	E2AB	E2BB	E2CB	E2DB	E2EB	E2FB	E30B	E31B	E32B
கௌ	ஙௌ	சௌ	ஞௌ	டௌ	ணௌ	தௌ	நௌ	பௌ	மௌ	யௌ	ரௌ	லௌ	வௌ	ழௌ	ளௌ	றௌ	னௌ
E21C	E22C	E23C	E24C	E25C	E26C	E27C	E28C	E29C	E2AC	E2BC	E2CC	E2DC	E2EC	E2FC	E30C	E31C	E32C

TABLE 3: TAMIL GRANTHA CONSONANTS

ஐ E330	ஸ E340	ஷ் E350	ஸ் E360	ஹ் E370	க்ஷ் E380
ஐ E331	ஸ E341	ஷ E351	ஸ E361	ஹ E371	க்ஷ E381
ஜா E332	ஸா E342	ஷா E352	ஸா E362	ஹா E372	க்ஷா E382
ஜி E333	ஸி E343	ஷி E353	ஸி E363	ஹி E373	க்ஷி E383
ஜீ E334	ஸீ E344	ஷீ E354	ஸீ E364	ஹீ E374	க்ஷீ E384
ஜு E335	ஸு E345	ஷு E355	ஸு E365	ஹு E375	க்ஷு E385
ஜூ E336	ஸூ E346	ஷூ E356	ஸூ E366	ஹூ E376	க்ஷூ E386
ஜெ E337	ஸெ E347	ஷெ E357	ஸெ E367	ஹெ E377	க்ஷெ E387
ஜே E338	ஸே E348	ஷே E358	ஸே E368	ஹே E378	க்ஷே E388
ஜை E339	ஸை E349	ஷை E359	ஸை E369	ஹை E379	க்ஷை E389
ஜொ E33A	ஸொ E34A	ஷொ E35A	ஸொ E36A	ஹொ E37A	க்ஷொ E38A
ஜோ E33B	ஸோ E34B	ஷோ E35B	ஸோ E36B	ஹோ E37B	க்ஷோ E38B
ஜௌ E33C	ஸௌ E34C	ஷௌ E35C	ஸௌ E36C	ஹௌ E37C	க்ஷௌ E38C
					ஸ்ரீ E38D