

Zeichen in Unicode für die elektronische Verarbeitung von Namen und den Datenaustausch in Europa

Latin characters in Unicode for the electronic processing of names and data exchange in Europe

Caractères latins en Unicode pour le traitement électronique des noms et l'échange de données en Europe

Dieser Entwurf mit Erscheinungsdatum 2018-08-03 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Spezifikation von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfs besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal von DIN unter www.din.de/go/entwuerfe bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an info@din.de möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/go/stellungnahmen-norm-entwuerfe oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an das DIN e. V., 10772 Berlin, Burggrafenstr. 6, 10787 Berlin.

Die Empfänger dieses Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Zur Erstellung einer DIN SPEC können verschiedene Verfahrensweisen herangezogen werden:
Das vorliegende Dokument wurde nach den Verfahrensregeln einer PAS erstellt.

Gesamtumfang 61 Seiten

Dieses Dokument wurde durch die im Vorwort genannten Verfasser erarbeitet und verabschiedet.

Inhalt

	Seite
Vorwort	3
Einleitung	5
1 Anwendungsbereich	8
2 Normative Verweisungen	8
3 Begriffe	9
4 Normative Schriftzeichen	10
4.1 Allgemeines	10
4.2 Lateinische Buchstaben	10
4.3 Kombinerende diakritische Zeichen	24
4.4 Nicht-Buchstaben N1	25
4.5 Nicht-Buchstaben N2	25
4.6 Nicht-Buchstaben N3	27
5 Normative Abbildung lateinischer Buchstaben auf Grundbuchstaben	27
6 Erweiterte Schriftzeichen (nicht-normativ)	41
6.1 Griechische Buchstaben	41
6.2 Kyrillische Buchstaben	43
6.3 Nicht-Buchstaben E1	45
7 Technische Datentypen (nicht-normativ)	46
7.1 Allgemeines	46
7.2 Datentyp A	47
7.3 Datentyp B	49
7.4 Datentyp C	51
7.5 Datentyp D	53
7.6 Datentyp E	55
Anhang A (informativ) Tabellarische Übersicht aller Zeichen	59
Anhang B (informativ) Mappings auf Legacy Codepages	60
Literaturhinweise	61

Vorwort

Diese DIN SPEC wurde nach dem PAS-Verfahren erarbeitet. Die Erarbeitung von DIN SPEC nach dem PAS-Verfahren erfolgt in Workshops und nicht zwingend unter Einbeziehung aller interessierten Kreise.

Die Erarbeitung und Verabschiedung des Dokuments erfolgten durch die nachfolgend genannten Initiatoren und Verfasser:

- Koordinierungsstelle für IT-Standards (KoSIT)
Frank Steimke
-]init[AG für digitale Kommunikation
Rico Aplitz
- Amtsgericht Essen
Helmut Lansing
- Anstalt für Kommunale Datenverarbeitung in Bayern
Franz-Xaver Salat
- Bayerisches Landesamt für Finanzen (LfF)
Martina Weiß
- Bayerisches Landesamt für Steuern, Projekt ELSTER
Franz Widholm
- Brandenburgisches Oberlandesgericht
Susann Lillienthal
- Bundesagentur für Arbeit
Alina Lorenz
- Bundesministerium des Innern (BMI), Referat IT I 3 - IT-Steuerung des Bundes, Standardisierung und Open Source
Oliver Tings
- Bundesministerium des Innern, für Bau und Heimat (BMI), Referat V II 1 – Verwaltungsverfahrenswesen, Informationszugangsrecht, Personenstandswesen
Rainer Bockstette
- Bundesverband der Deutschen Standesbeamtinnen und Standesbeamten e.V. (BDS), Akademie für Personenstandswesen, Akademie für Staatsangehörigkeitsrecht und Meldewesen
Beate Tripp
- Deutsche Rentenversicherung Bund, Bereich 0500
Werner Meckelein, Dr. Clemens Mühlberger,
- Deutsche Rentenversicherung Bund, Zentrale Zulagenstelle für Altersvermögen (ZfA)
Silke Klostermann
- Freelancer
Bernd Kappenberg

- Landeshauptstadt München, Referat für Informations- und Telekommunikationstechnik, Hauptabteilung I - IT-Strategie und IT-Steuerung / IT-Controlling (STRAC)
Dr. Marlen Jurisch
- Westernacher Solutions AG
Volker Dittmar

Für dieses Thema bestehen derzeit keine Normen im Deutschen Normenwerk.

DIN SPEC sind nicht Teil des Deutschen Normenwerks.

Es handelt sich bei dieser Veröffentlichung um den Entwurf der DIN SPEC zur Kommentierung durch die Öffentlichkeit.

Trotz großer Anstrengungen zur Sicherstellung der Korrektheit, Verlässlichkeit und Präzision technischer und nicht-technischer Beschreibungen kann der Workshop weder eine explizite noch eine implizite Gewährleistung für die Korrektheit des Dokuments übernehmen. Die Benutzung dieses Dokuments geschieht in dem Bewusstsein, dass der Workshop für Schäden oder Verluste jeglicher Art nicht haftbar gemacht werden kann. Die Anwendung der vorliegenden DIN SPEC entbindet den Nutzer nicht von der Verantwortung für eigenes Handeln und geschieht damit auf eigene Gefahr.

Die kostenfreie Bereitstellung dieses Dokuments als PDF-Version über den Beuth Webshop wurde im Vorfeld finanziert.

Einleitung

Ausgangslage

Buchstaben und andere Schriftzeichen (Nicht-Buchstaben) können in IT-Verfahren nur verarbeitet werden, indem sie auf Zahlen abgebildet werden. Es gibt auf der Welt viele verschiedene Alphabete. Für die meisten davon gibt es wiederum viele unterschiedliche Abbildungen der enthaltenen Schriftzeichen auf Zahlen (Zeichenkodierungen oder Encodings). Aus dem Nebeneinander verschiedener Kodierungssysteme für unterschiedliche Alphabete können Interoperabilitätsprobleme resultieren. Der Unicode-Standard bzw. die Norm ISO/IEC 10646 wurde entwickelt, um diese Probleme zu lösen.

Die Forderung nach vollständiger Umsetzung des Unicode-Standards auch auf der Anwendungsebene von IT-Verfahren ist jedoch in der Regel weder sachgerecht noch wirtschaftlich, weil der kulturelle Kontext der Anwender und des intendierten Anwendungsbereichs zu berücksichtigen sind.

IT-Verfahren der öffentlichen Verwaltung Deutschlands sind regelhaft auf die Verarbeitung von Daten in deutscher Sprache ausgerichtet, denn die Amtssprache ist Deutsch. Werden bei einer Behörde in einer fremden Sprache Anträge gestellt oder Eingaben, Belege, Urkunden oder sonstige Dokumente vorgelegt, soll die Behörde unverzüglich die Vorlage einer Übersetzung verlangen (siehe beispielsweise § 23 Verwaltungsverfahrensgesetz). Die deutsche Sprache nutzt die lateinische Schrift. Sofern Dokumente oder Urkunden vorgelegt werden, die Angaben in anderen als lateinischen Schriftzeichen wiedergibt, dann sollen diese in die lateinischen Schriftzeichen transliteriert werden, bevor die Angaben in IT-Verfahren und automatisiert geführte Register übernommen werden. Dies gilt insbesondere für die Namen natürlicher Personen. Entsprechend des Übereinkommens über die Angabe von Familiennamen und Vornamen in den Personenstandsbüchern vom 13. September 1973 (BGBl. 1976 II, Seite 1473 f.) sind diese ohne Übersetzung wiederzugeben bzw. in IT-Verfahren und Register zu übernehmen. Sofern die Namen in lateinischen Schriftzeichen vorliegen, sind sie buchstabengetreu ohne Änderung wiederzugeben. Diakritische Zeichen sind ebenfalls unverändert wiederzugeben, selbst wenn die deutsche Sprache diese Zeichen nicht kennt. Werden Namen hingegen in anderen als lateinischen Schriftzeichen vorgelegt bzw. übermittelt, so sind sie soweit wie möglich durch Transliteration wiederzugeben. Hierfür sollen die von der Internationalen Normenorganisation (ISO) empfohlenen Normen genutzt werden, siehe Literaturhinweise. Diese Vorgaben gelten analog auch für elektronische Datenübermittlungen.

Standardisierungsbedarf

Aus der Ausgangslage ergibt sich die Notwendigkeit, für IT-Verfahren (insbesondere solche der öffentlichen Verwaltung Deutschlands)

- Einerseits festzulegen, dass alle Schriftzeichen der lateinischen Schrift, die auf hoheitlichen Dokumenten und Urkunden verwendet werden können, vollumfänglich unterstützt werden müssen;
- Andererseits die Möglichkeit zu eröffnen, andere Schriftzeichen von Datenübermittlungen auszuschließen, da diese in den meisten IT-gestützten Prozessen der deutschen Verwaltung nicht akzeptiert werden sollen.

Basierend auf dem Standard Unicode ist somit die Menge der darin enthaltenen, für das Verwaltungshandeln notwendigen und zulässigen Schriftzeichen festzulegen. Diese Teilmenge von Unicode muss anhand fachlicher Überlegungen bestimmt werden, zu denen insbesondere die folgenden gehören:

- a) Welche lateinischen Schriftzeichen werden bei der Erfassung und Nacherfassung von Personenstandsdaten für die elektronisch geführten Personenstandsregister festgestellt?
- b) Welche lateinischen Schriftzeichen können als Ergebnis einer Transliteration (d.h. der buchstabengetreuen Übertragung von Wörtern aus einer Schrift in eine andere, sodass eine eindeutige Rückübertragung möglich wird) aus anderen Schriften auftreten? Dabei müssen die Transliterationsregeln der ISO zugrunde gelegt werden.

- c) Welche lateinischen Schriftzeichen werden in den europäischen Amtssprachen genutzt?
- d) Welche lateinischen Schriftzeichen werden für den Datenaustausch zwischen Sozialversicherungsträgern benötigt?
- e) Welche Schriftzeichen werden für Einträge im Handelsregister benötigt?
- f) Welche Schriftzeichen werden für den EU-weiten Datenaustausch inkl. Island, Liechtenstein, Norwegen und der Schweiz zwischen Behörden benötigt?

Abwägung von Alternativen

Die vollständige Unterstützung aller Unicode-Blöcke zu fordern, in denen unter anderem auch lateinische Zeichen enthalten sind, ist keine Option, denn:

- Bei der Bestimmung von Zeichenblöcken durch das Unicode-Konsortium spielen die oben genannten, fachlichen Aspekte keine oder nur eine untergeordnete Rolle;
- Es sind auch historische Zeichen ohne Relevanz für das aktuelle Verwaltungshandeln enthalten. Es wäre weder sachgerecht noch wirtschaftlich, deren technische Umsetzung zu fordern;
- Für Transliterationen notwendige kodierte Zeichensequenzen wären nicht explizit im Standard aufgeführt;
- Die Eingabe von ungewöhnlichen Schriftzeichen ist schwieriger und fehleranfälliger, wenn mehr Schriftzeichen unterstützt werden müssen, als man braucht;
- Eine abschließende Liste der zulässigen Schriftzeichen erleichtert die Implementation robuster Algorithmen und Verfahren, um Personen zu identifizieren und hoheitliche Dokumente und Urkunden zu erstellen.

Außerdem gibt es viele Anwendungsfälle, wie zum Beispiel die Ausstellung von Ausweisen, Pässen und anderen hoheitlichen Dokumenten, in denen die Verwendung zusätzlicher Schriftzeichen ein Sicherheitsrisiko darstellen kann. Denn die Freiheit von Unicode, Grundbuchstaben beliebig mit Diakritika zu kombinieren, schafft eine Möglichkeit, Schriftzeichen zu erzeugen, die anderen Schriftzeichen zum Verwechseln ähnlich sehen – IT-technisch aber verschieden sind. Böswillig konstruierte kodierte Zeichensequenzen können verwendet werden, um Identitäten zu verschleiern.

Zudem könnte sich bei der hier angewandten (informations)technischen Sicht (d.h. die Zeichenlänge entspricht der Anzahl der Codepoints) die Zeichenlänge durch Zeichenkombinationen erhöhen: beispielsweise ließe sich das Zeichen „ä“ durch einen Codepoint (00E4) oder als Kombination der beiden Codepoints 0061 (Zeichen „a“) und 0308 (kombinierendes diakritisches Zeichen „“) darstellen.

Durch die DIN SPEC lassen sich diese Risiken ausschließen oder zumindest erheblich reduzieren. Denn sie kann in fachlich begründeten Fällen für die Einschränkung genutzt werden, ausschließlich die darin genannten Zeichen zu verarbeiten bzw. zu übermitteln. Die Durchsetzung einer entsprechenden Einschränkung kann sowohl rechtlich als auch technisch erfolgen.

Die genannten Gründe führen zu dem Ergebnis, dass eine abschließende Auflistung der zu unterstützenden Schriftzeichen als eine echte Teilmenge der in Unicode definierten Zeichen in der DIN SPEC notwendig ist.

Überblick

Im Sinne des Investitionsschutzes berücksichtigt die DIN SPEC den vom IT-Planungsrat und der Innenministerkonferenz derzeit bereits für den gleichen Anwendungsbereich verbindlich vorgegebenen Standard „Lateinische Zeichen in Unicode, Version 1.1 vom 27.1.2012“ angemessen. Die vorliegende DIN SPEC enthält Zeichen und Zeichensequenzen in folgenden Gruppen (siehe Tabelle 1):

Tabelle 1 — Gruppen von Zeichen und Zeichensequenzen

Gruppe	Summe
Latein. Buchstaben (normativ) (char)	494
Latein. Buchstaben (normativ) (sequence)	143
Komb. diakrit. Zeichen (normativ) (char)	19
Nicht-Buchstaben N1 (normativ) (char)	16
Nicht-Buchstaben N2 (normativ) (char)	60
Nicht-Buchstaben N3 (normativ) (char)	4
Griech. Buchstaben (erweitert) (char)	69
Kyrill. Buchstaben (erweitert) (char)	60
Nicht-Buchstaben E1 (erweitert) (char)	46
Insgesamt	911

Das „char!“ steht für Zeichen mit einem Codepoint und das "sequence" steht für Zeichensequenzen bestehend aus mehreren Codepoints.

1 Anwendungsbereich

Diese DIN SPEC 91379 „Zeichen in Unicode für die elektronische Verarbeitung von Namen und den Datenaustausch in Europa“ legt die Zeichen fest, die für Namen im weiteren Sinne benötigt werden. Namen im weiteren Sinne sind Namen natürlicher Personen gemäß der Vorgaben des Personenstandsrechts ebenso wie Namen juristischer Personen, Namen von Produkten, Orten und Straßen sowie Titel von Dokumenten oder Gesetzen. Namen im weiteren Sinne können konkrete Objekte bezeichnen, aber auch virtuelle Konstrukte wie Produktgruppen oder Musikstile.

Die DIN SPEC teilt sich auf in einen normativen Teil und einen nicht-normativen Teil.

Der normative Teil der DIN SPEC legt diejenige Teilmenge der in Unicode enthaltenen Zeichen und Zeichensequenzen fest, die für IT-Verfahren für die elektronische Verarbeitung von Namen auf Basis der lateinischen Schrift erforderlich ist und deshalb von allen zu dieser Spezifikation konformen IT-Verfahren unterstützt werden muss. In diesem Teil wird auch eine gegebenenfalls erforderliche Abbildung der normativen Buchstaben auf die Großbuchstaben A bis Z bestimmt.

Im nicht-normativen Teil der DIN SPEC werden darüber hinaus ergänzende Zeichen aus Unicode festgelegt, die eventuell zur Bildung von Namen juristischer Personen oder von Produkten benötigt werden. Im Kontext des EU-weiten Datenaustauschs ergibt sich ein erweiterter Bedarf an griechischen und kyrillischen Zeichen. In diesem Teil werden auch reguläre Ausdrücke für technische Datentypen vorgeschlagen.

Durch den normativen Teil wird eine Mindestmenge von Schriftzeichen bestimmt. Zur DIN SPEC konforme IT-Verfahren dürfen darüber hinausgehende Schriftzeichen unterstützen. In Schnittstellenvereinbarungen ist es jedoch erforderlich, den jeweiligen Zeichenvorrat abschließend festzulegen. Die DIN SPEC stellt in Abschnitt 7 reguläre Ausdrücke für fachlich bestimmte Anwendungsfälle zur Verfügung. Dazu werden bestimmte Schriftzeichengruppen zu einer Menge zulässiger Schriftzeichen zusammengestellt. Die im nicht-normativen Teil empfohlenen regulären Ausdrücke akzeptieren ausschließlich die jeweils zugrunde liegenden Schriftzeichen und können in technischen Datentypen implementiert werden.

Diese DIN SPEC basiert auf dem Unicode-Standard und wendet sich vorrangig an Behörden und Organisationen, die IT-Verfahren betreiben, welche dem behördenübergreifenden Datenaustausch oder dem Datenaustausch mit Bürgern und Wirtschaft dienen. Sie deckt in ihrer Gesamtheit die EU-Amtssprachen und die Amtssprachen Islands, Liechtensteins, Norwegens und der Schweiz sowie die deutschen Minderheitensprachen ab.

Alle in dieser DIN SPEC genannten Zeichen und Zeichensequenzen liegen in der *Normalization Form C (NFC)* des *Unicode Standard Annex #15 „Unicode Normalization Forms“* vor.

Die DIN SPEC trifft weder Aussagen zum Umgang mit historischen Zeichen noch zum Umgang mit Fließtext. Die DIN SPEC regelt nicht die Darstellung von Zeichen (Glyphen). Die in dieser DIN SPEC verwendeten Glyphen orientieren sich an der für die ISO/IEC 10646 gewählten Darstellung

2 Normative Verweisungen

Die folgenden Dokumente werden im Text in solcher Weise in Bezug genommen, dass einige Teile davon oder ihr gesamter Inhalt Anforderungen des vorliegenden Dokuments darstellen. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*

The Unicode Standard, *Version 9.0 — Core Specification*¹⁾

1) Available at: <https://www.unicode.org/versions/Unicode9.0.0>

3 Begriffe

Die für diese Spezifikation relevanten Definitionen des Unicode-Standards werden unverändert übernommen. Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

3.1

abstraktes Zeichen

(en: abstract character)

Informationseinheit, die für die Organisation, Kontrolle oder Darstellung von Textdaten verwendet wird

[QUELLE: The Unicode Standard, Version 9.0, 3.4, D.7]

3.2

Codepoint

Wert innerhalb des Wertebereichs der ganzen Zahlen von 0 bis $10FFFF_{16}$

Anmerkung 1 zum Begriff: Der Wert repräsentiert die Position eines Zeichens im Unicode-Standard.

Anmerkung 2 zum Begriff: Für diese DIN SPEC werden nur Werte bis zu $FFFF_{16}$ verwendet, sodass jeder Codepoint binär mit zwei Byte dargestellt werden kann.

[QUELLE: The Unicode Standard, Version 9.0, 3.4, D.10 in Verbindung mit D.9]

3.3

kodiertes Zeichen

(en: encoded character, coded character)

Zeichen

(en: character)

Verbindung (oder Abbildung) zwischen einem abstrakten Zeichen und einem Codepoint

Anmerkung 1 zum Begriff: Der Begriff „Zeichen“ ist allein stehend als Abkürzung für den Begriff „kodiertes Zeichen“ zu verstehen. Siehe hierfür die Erläuterung in Unicode 9.0, 3.4, zwischen den Definitionen D.12 und D.13.

[QUELLE: The Unicode Standard, Version 9.0, 3.4, D.11]

3.4

kodierte Zeichensequenz

(en: coded character sequence)

Zeichensequenz

(en: character sequence)

geordnete Sequenz eines oder mehrerer Codepoints

Anmerkung 1 zum Begriff: Der Begriff „Zeichensequenz“ steht als Abkürzung für den Begriff „kodierte Zeichensequenz“. Siehe hierfür die Erläuterung in Unicode 9.0, 3.4, zwischen den Definitionen D.12 und D.13.

[QUELLE: The Unicode Standard, Version 9.0, 3.4, D.12]

3.5

Schriftzeichen

kleinste im Schreibfluss aufeinanderfolgende Einheiten einer Schrift, z. B. des lateinischen oder kyrillischen Alphabets, die sprachunabhängig sind, da sie für verschiedene Sprachen genutzt werden können, z. B. das lateinische Alphabet für Deutsch und Französisch

Anmerkung 1 zum Begriff: Zu den Schriftzeichen gehören sowohl kodierte Zeichen als auch kodierte Zeichensequenzen, jedoch keine einzelnen kombinierenden diakritischen Zeichen außerhalb von Zeichensequenzen.

Anmerkung 2 zum Begriff: Im Rahmen dieser DIN SPEC werden die Schriftzeichen unterschieden in Buchstaben und Nicht-Buchstaben (z. B. Interpunktion, Ziffern und Leerzeichen).

Anmerkung 3 zum Begriff: Die Menge der Schriftzeichen, deren Unterstützung durch IT-Verfahren für Zwecke der elektronischen Verarbeitung von Namen und den Datenaustausch in Europa vorgeschrieben bzw. empfohlen wird, ist eine echte Teilmenge der im Unicode-Standard insgesamt enthaltenen Schriftzeichen.

Anmerkung 4 zum Begriff: In dieser DIN SPEC wird der Begriff „Schriftzeichen“ weitgehend synonym zum Begriff „Graphem“ verwendet. Es werden aber auch „Plurigraphie“ mit eingeschlossen.

3.6

Schriftzeichengruppe

Zusammenfassung von Schriftzeichen unter fachlichen Aspekten

Anmerkung 1 zum Begriff: Jedes der in dieser Spezifikation genannten Schriftzeichen ist genau einer Schriftzeichengruppe zugeordnet. Jeder Schriftzeichengruppe ist mindestens ein Schriftzeichen zugeordnet. Jede Schriftzeichengruppe hat einen innerhalb dieser Spezifikation eindeutigen Namen, z. B. "Lateinische Buchstaben" oder "Nicht-Buchstaben N1".

4 Normative Schriftzeichen

4.1 Allgemeines

Jedes konforme IT-Verfahren MUSS die in Tabelle 2 bis Tabelle 6 genannten Schriftzeichen bei allen für Namen (im weiteren Sinne) bestimmten Datenfeldern in vollem Umfang verarbeiten können, weil sie für die Darstellung von Namen im weiteren Sinne und für die Interoperabilität zu ISO 8859-15 zwingend erforderlich sind. Dies umfasst insbesondere die Erfassung, Speicherung, Übermittlung, Anzeige und den Ausdruck aller Schriftzeichen.

4.2 Lateinische Buchstaben

Die in Tabelle 2 genannten Schriftzeichen für lateinische Buchstaben sind für die Darstellung von Namen im weiteren Sinne zwingend erforderlich.

Tabelle 2 — Normative lateinische Buchstaben

Nr.	Codepoints	Name	Char
001	U+0041	LATIN CAPITAL LETTER A	A
002	<U+0041,U+030B>	LATIN CAPITAL LETTER A WITH COMBINING DOUBLE ACUTE ACCENT	Ă
003	U+0042	LATIN CAPITAL LETTER B	B
004	U+0043	LATIN CAPITAL LETTER C	C
005	<U+0043,U+0300>	LATIN CAPITAL LETTER C WITH COMBINING GRAVE ACCENT	Ĉ
006	<U+0043,U+0304>	LATIN CAPITAL LETTER C WITH COMBINING MACRON	Ċ
007	<U+0043,U+0306>	LATIN CAPITAL LETTER C WITH COMBINING BREVE	Č
008	<U+0043,U+0308>	LATIN CAPITAL LETTER C WITH COMBINING DIAERESIS	Č
009	<U+0043,U+0315>	LATIN CAPITAL LETTER C WITH COMBINING COMMA ABOVE RIGHT	Ć
010	<U+0043,U+0323>	LATIN CAPITAL LETTER C WITH COMBINING DOT BELOW	Ç
011	<U+0043,U+0326>	LATIN CAPITAL LETTER C WITH COMBINING COMMA BELOW	Ç
012	<U+0043,U+0328,U+0306>	LATIN CAPITAL LETTER C WITH COMBINING OGONEK AND COMBINING BREVE	Ĉ
013	U+0044	LATIN CAPITAL LETTER D	D
014	<U+0044,U+0302>	LATIN CAPITAL LETTER D WITH COMBINING CIRCUMFLEX ACCENT	Ď
015	U+0045	LATIN CAPITAL LETTER E	E
016	U+0046	LATIN CAPITAL LETTER F	F
017	<U+0046,U+0300>	LATIN CAPITAL LETTER F WITH COMBINING GRAVE ACCENT	Ĕ
018	<U+0046,U+0304>	LATIN CAPITAL LETTER F WITH COMBINING MACRON	Ė
019	U+0047	LATIN CAPITAL LETTER G	G

Nr.	Codepoints	Name	Char
020	<U+0047,U+0300>	LATIN CAPITAL LETTER G WITH COMBINING GRAVE ACCENT	Ġ
021	U+0048	LATIN CAPITAL LETTER H	H
022	<U+0048,U+0304>	LATIN CAPITAL LETTER H WITH COMBINING MACRON	Ĥ
023	<U+0048,U+0326>	LATIN CAPITAL LETTER H WITH COMBINING COMMA BELOW	Ħ
024	<U+0048,U+0331>	LATIN CAPITAL LETTER H WITH COMBINING MACRON BELOW	Ħ̃
025	U+0049	LATIN CAPITAL LETTER I	I
026	U+004A	LATIN CAPITAL LETTER J	J
027	<U+004A,U+0301>	LATIN CAPITAL LETTER J WITH COMBINING ACUTE ACCENT	Ĵ
028	<U+004A,U+030C>	LATIN CAPITAL LETTER J WITH COMBINING CARON	Ĵ̃
029	U+004B	LATIN CAPITAL LETTER K	K
030	<U+004B,U+0300>	LATIN CAPITAL LETTER K WITH COMBINING GRAVE ACCENT	Ķ
031	<U+004B,U+0302>	LATIN CAPITAL LETTER K WITH COMBINING CIRCUMFLEX ACCENT	Ķ̂
032	<U+004B,U+0304>	LATIN CAPITAL LETTER K WITH COMBINING MACRON	Ķ̄
033	<U+004B,U+0307>	LATIN CAPITAL LETTER K WITH COMBINING DOT ABOVE	Ķ̇
034	<U+004B,U+0315>	LATIN CAPITAL LETTER K WITH COMBINING COMMA ABOVE RIGHT	Ķ̌
035	<U+004B,U+031B>	LATIN CAPITAL LETTER K WITH COMBINING HORN	Ķ̊
036	<U+004B,U+0326>	LATIN CAPITAL LETTER K WITH COMBINING COMMA BELOW	Ķ̸
037	<U+004B,U+035F,U+0048>	LATIN CAPITAL LETTER K WITH COMBINING DOUBLE MACRON BELOW AND LATIN CAPITAL LETTER H	Ķ̹
038	<U+004B,U+035F,U+0068>	LATIN CAPITAL LETTER K WITH COMBINING DOUBLE MACRON BELOW AND LATIN SMALL LETTER H	Ķ̹̈
039	U+004C	LATIN CAPITAL LETTER L	L
040	<U+004C,U+0302>	LATIN CAPITAL LETTER L WITH COMBINING CIRCUMFLEX ACCENT	Ĺ
041	<U+004C,U+0325>	LATIN CAPITAL LETTER L WITH COMBINING RING BELOW	Ļ
042	<U+004C,U+0325,U+0304>	LATIN CAPITAL LETTER L WITH COMBINING RING BELOW AND COMBINING MACRON	Ļ̄
043	<U+004C,U+0326>	LATIN CAPITAL LETTER L WITH COMBINING COMMA BELOW	Ļ̸
044	U+004D	LATIN CAPITAL LETTER M	M
045	<U+004D,U+0300>	LATIN CAPITAL LETTER M WITH COMBINING GRAVE ACCENT	Ĺ̂
046	<U+004D,U+0302>	LATIN CAPITAL LETTER M WITH COMBINING CIRCUMFLEX ACCENT	Ĺ̂̂
047	<U+004D,U+0306>	LATIN CAPITAL LETTER M WITH COMBINING BREVE	Ĺ̆
048	<U+004D,U+0310>	LATIN CAPITAL LETTER M WITH COMBINING CANDRABINDU	Ĺ̇
049	U+004E	LATIN CAPITAL LETTER N	N
050	<U+004E,U+0302>	LATIN CAPITAL LETTER N WITH COMBINING CIRCUMFLEX ACCENT	Ñ
051	<U+004E,U+0304>	LATIN CAPITAL LETTER N WITH COMBINING MACRON	Ñ̄
052	<U+004E,U+0306>	LATIN CAPITAL LETTER N WITH COMBINING BREVE	Ñ̆
053	<U+004E,U+0326>	LATIN CAPITAL LETTER N WITH COMBINING COMMA BELOW	Ñ̸
054	U+004F	LATIN CAPITAL LETTER O	O
055	U+0050	LATIN CAPITAL LETTER P	P
056	<U+0050,U+0300>	LATIN CAPITAL LETTER P WITH COMBINING GRAVE ACCENT	Ĵ̂
057	<U+0050,U+0304>	LATIN CAPITAL LETTER P WITH COMBINING MACRON	Ĵ̄
058	<U+0050,U+0315>	LATIN CAPITAL LETTER P WITH COMBINING COMMA ABOVE RIGHT	Ĵ̌
059	<U+0050,U+0323>	LATIN CAPITAL LETTER P WITH COMBINING DOT BELOW	Ĵ̇
060	U+0051	LATIN CAPITAL LETTER Q	Q
061	U+0052	LATIN CAPITAL LETTER R	R
062	<U+0052,U+0306>	LATIN CAPITAL LETTER R WITH COMBINING BREVE	Ŕ
063	<U+0052,U+0325>	LATIN CAPITAL LETTER R WITH COMBINING RING BELOW	Ŗ

Nr.	Codepoints	Name	Char
064	<U+0052,U+0325,U+0304>	LATIN CAPITAL LETTER R WITH COMBINING RING BELOW AND COMBINING MACRON	Ř
065	U+0053	LATIN CAPITAL LETTER S	S
066	<U+0053,U+0300>	LATIN CAPITAL LETTER S WITH COMBINING GRAVE ACCENT	Š
067	<U+0053,U+0304>	LATIN CAPITAL LETTER S WITH COMBINING MACRON	Ŝ
068	<U+0053,U+031B,U+0304>	LATIN CAPITAL LETTER S WITH COMBINING HORN AND COMBINING MACRON	Š̄
069	<U+0053,U+0331>	LATIN CAPITAL LETTER S WITH COMBINING MACRON BELOW	Ŧ
070	U+0054	LATIN CAPITAL LETTER T	T
071	<U+0054,U+0300>	LATIN CAPITAL LETTER T WITH COMBINING GRAVE ACCENT	Ť
072	<U+0054,U+0304>	LATIN CAPITAL LETTER T WITH COMBINING MACRON	Ť̄
073	<U+0054,U+0308>	LATIN CAPITAL LETTER T WITH COMBINING DIAERESIS	Ŧ̈
074	<U+0054,U+0315>	LATIN CAPITAL LETTER T WITH COMBINING COMMA ABOVE RIGHT	Ŧ̸
075	<U+0054,U+031B>	LATIN CAPITAL LETTER T WITH COMBINING HORN	Ŧ̸̈
076	U+0055	LATIN CAPITAL LETTER U	U
077	<U+0055,U+0307>	LATIN CAPITAL LETTER U WITH COMBINING DOT ABOVE	Û
078	U+0056	LATIN CAPITAL LETTER V	V
079	U+0057	LATIN CAPITAL LETTER W	W
080	U+0058	LATIN CAPITAL LETTER X	X
081	U+0059	LATIN CAPITAL LETTER Y	Y
082	U+005A	LATIN CAPITAL LETTER Z	Z
083	<U+005A,U+0304>	LATIN CAPITAL LETTER Z WITH COMBINING MACRON	Ž
084	<U+005A,U+0306>	LATIN CAPITAL LETTER Z WITH COMBINING BREVE	Ž̆
085	<U+005A,U+0308>	LATIN CAPITAL LETTER Z WITH COMBINING DIAERESIS	Ž̈
086	<U+005A,U+0327>	LATIN CAPITAL LETTER Z WITH COMBINING CEDILLA	Ʒ
087	U+0061	LATIN SMALL LETTER A	a
088	<U+0061,U+030B>	LATIN SMALL LETTER A WITH COMBINING DOUBLE ACUTE ACCENT	ă
089	U+0062	LATIN SMALL LETTER B	b
090	U+0063	LATIN SMALL LETTER C	c
091	<U+0063,U+0300>	LATIN SMALL LETTER C WITH COMBINING GRAVE ACCENT	ç
092	<U+0063,U+0304>	LATIN SMALL LETTER C WITH COMBINING MACRON	ċ
093	<U+0063,U+0306>	LATIN SMALL LETTER C WITH COMBINING BREVE	ċ̆
094	<U+0063,U+0308>	LATIN SMALL LETTER C WITH COMBINING DIAERESIS	ċ̈
095	<U+0063,U+0315>	LATIN SMALL LETTER C WITH COMBINING COMMA ABOVE RIGHT	ć̸
096	<U+0063,U+0323>	LATIN SMALL LETTER C WITH COMBINING DOT BELOW	ċ̣
097	<U+0063,U+0326>	LATIN SMALL LETTER C WITH COMBINING COMMA BELOW	ċ̥
098	<U+0063,U+0328,U+0306>	LATIN SMALL LETTER C WITH COMBINING OGONEK AND COMBINING BREVE	ċ̆̈
099	U+0064	LATIN SMALL LETTER D	d
100	<U+0064,U+0302>	LATIN SMALL LETTER D WITH COMBINING CIRCUMFLEX ACCENT	đ
101	U+0065	LATIN SMALL LETTER E	e
102	U+0066	LATIN SMALL LETTER F	f
103	<U+0066,U+0300>	LATIN SMALL LETTER F WITH COMBINING GRAVE ACCENT	ƒ
104	<U+0066,U+0304>	LATIN SMALL LETTER F WITH COMBINING MACRON	ƒ̄
105	U+0067	LATIN SMALL LETTER G	g
106	<U+0067,U+0300>	LATIN SMALL LETTER G WITH COMBINING GRAVE ACCENT	ğ
107	U+0068	LATIN SMALL LETTER H	h

Nr.	Codepoints	Name	Char
108	<U+0068,U+0304>	LATIN SMALL LETTER H WITH COMBINING MACRON	ĥ
109	<U+0068,U+0326>	LATIN SMALL LETTER H WITH COMBINING COMMA BELOW	ħ
110	U+0069	LATIN SMALL LETTER I	i
111	U+006A	LATIN SMALL LETTER J	j
112	<U+006A,U+0301>	LATIN SMALL LETTER J WITH COMBINING ACUTE ACCENT	ĵ
113	U+006B	LATIN SMALL LETTER K	k
114	<U+006B,U+0300>	LATIN SMALL LETTER K WITH COMBINING GRAVE ACCENT	ķ
115	<U+006B,U+0302>	LATIN SMALL LETTER K WITH COMBINING CIRCUMFLEX ACCENT	ķ
116	<U+006B,U+0304>	LATIN SMALL LETTER K WITH COMBINING MACRON	ķ̄
117	<U+006B,U+0307>	LATIN SMALL LETTER K WITH COMBINING DOT ABOVE	ķ̇
118	<U+006B,U+0315>	LATIN SMALL LETTER K WITH COMBINING COMMA ABOVE RIGHT	ķ'
119	<U+006B,U+031B>	LATIN SMALL LETTER K WITH COMBINING HORN	ķ̊
120	<U+006B,U+0326>	LATIN SMALL LETTER K WITH COMBINING COMMA BELOW	ķ̸
121	<U+006B,U+035F,U+0068>	LATIN SMALL LETTER K WITH COMBINING DOUBLE MACRON BELOW AND LATIN SMALL LETTER H	kh̶
122	U+006C	LATIN SMALL LETTER L	l
123	<U+006C,U+0302>	LATIN SMALL LETTER L WITH COMBINING CIRCUMFLEX ACCENT	Ĺ
124	<U+006C,U+0325>	LATIN SMALL LETTER L WITH COMBINING RING BELOW	ł̣
125	<U+006C,U+0325,U+0304>	LATIN SMALL LETTER L WITH COMBINING RING BELOW AND COMBINING MACRON	Ł̣̄
126	<U+006C,U+0326>	LATIN SMALL LETTER L WITH COMBINING COMMA BELOW	ł̸
127	U+006D	LATIN SMALL LETTER M	m
128	<U+006D,U+0300>	LATIN SMALL LETTER M WITH COMBINING GRAVE ACCENT	ṁ
129	<U+006D,U+0302>	LATIN SMALL LETTER M WITH COMBINING CIRCUMFLEX ACCENT	ṁ
130	<U+006D,U+0306>	LATIN SMALL LETTER M WITH COMBINING BREVE	ṁ
131	<U+006D,U+0310>	LATIN SMALL LETTER M WITH COMBINING CANDRABINDU	ṃ̇
132	U+006E	LATIN SMALL LETTER N	n
133	<U+006E,U+0302>	LATIN SMALL LETTER N WITH COMBINING CIRCUMFLEX ACCENT	ñ
134	<U+006E,U+0304>	LATIN SMALL LETTER N WITH COMBINING MACRON	ñ̄
135	<U+006E,U+0306>	LATIN SMALL LETTER N WITH COMBINING BREVE	ñ̆
136	<U+006E,U+0326>	LATIN SMALL LETTER N WITH COMBINING COMMA BELOW	ñ̸
137	U+006F	LATIN SMALL LETTER O	o
138	U+0070	LATIN SMALL LETTER P	p
139	<U+0070,U+0300>	LATIN SMALL LETTER P WITH COMBINING GRAVE ACCENT	ṑ
140	<U+0070,U+0304>	LATIN SMALL LETTER P WITH COMBINING MACRON	ṑ̄
141	<U+0070,U+0315>	LATIN SMALL LETTER P WITH COMBINING COMMA ABOVE RIGHT	ṑ'
142	<U+0070,U+0323>	LATIN SMALL LETTER P WITH COMBINING DOT BELOW	ọ̄̀
143	U+0071	LATIN SMALL LETTER Q	q
144	U+0072	LATIN SMALL LETTER R	r
145	<U+0072,U+0306>	LATIN SMALL LETTER R WITH COMBINING BREVE	ṛ
146	<U+0072,U+0325>	LATIN SMALL LETTER R WITH COMBINING RING BELOW	ṛ̣
147	<U+0072,U+0325,U+0304>	LATIN SMALL LETTER R WITH COMBINING RING BELOW AND COMBINING MACRON	Ṝ̣
148	U+0073	LATIN SMALL LETTER S	s
149	<U+0073,U+0300>	LATIN SMALL LETTER S WITH COMBINING GRAVE ACCENT	ṣ
150	<U+0073,U+0304>	LATIN SMALL LETTER S WITH COMBINING MACRON	ṣ̄
151	<U+0073,U+031B,U+0304>	LATIN SMALL LETTER S WITH COMBINING HORN AND COMBINING MACRON	ṣ̊̄

Nr.	Codepoints	Name	Char
152	<U+0073,U+0331>	LATIN SMALL LETTER S WITH COMBINING MACRON BELOW	ſ̶
153	U+0074	LATIN SMALL LETTER T	t
154	<U+0074,U+0300>	LATIN SMALL LETTER T WITH COMBINING GRAVE ACCENT	ṭ
155	<U+0074,U+0304>	LATIN SMALL LETTER T WITH COMBINING MACRON	ṫ
156	<U+0074,U+0315>	LATIN SMALL LETTER T WITH COMBINING COMMA ABOVE RIGHT	ṭ́
157	<U+0074,U+031B>	LATIN SMALL LETTER T WITH COMBINING HORN	ṭ̸
158	U+0075	LATIN SMALL LETTER U	u
159	<U+0075,U+0307>	LATIN SMALL LETTER U WITH COMBINING DOT ABOVE	ù
160	U+0076	LATIN SMALL LETTER V	v
161	U+0077	LATIN SMALL LETTER W	w
162	U+0078	LATIN SMALL LETTER X	x
163	U+0079	LATIN SMALL LETTER Y	y
164	U+007A	LATIN SMALL LETTER Z	z
165	<U+007A,U+0304>	LATIN SMALL LETTER Z WITH COMBINING MACRON	ž
166	<U+007A,U+0306>	LATIN SMALL LETTER Z WITH COMBINING BREVE	ẏ
167	<U+007A,U+0308>	LATIN SMALL LETTER Z WITH COMBINING DIAERESIS	ẑ
168	<U+007A,U+0327>	LATIN SMALL LETTER Z WITH COMBINING CEDILLA	Ẓ
169	U+00C0	LATIN CAPITAL LETTER A WITH GRAVE	À
170	U+00C1	LATIN CAPITAL LETTER A WITH ACUTE	Á
171	U+00C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX	Â
172	U+00C3	LATIN CAPITAL LETTER A WITH TILDE	Ã
173	U+00C4	LATIN CAPITAL LETTER A WITH DIAERESIS	Ä
174	U+00C5	LATIN CAPITAL LETTER A WITH RING ABOVE	Å
175	U+00C6	LATIN CAPITAL LETTER AE	Æ
176	U+00C7	LATIN CAPITAL LETTER C WITH CEDILLA	Ç
177	U+00C8	LATIN CAPITAL LETTER E WITH GRAVE	È
178	U+00C9	LATIN CAPITAL LETTER E WITH ACUTE	É
179	U+00CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX	Ê
180	U+00CB	LATIN CAPITAL LETTER E WITH DIAERESIS	Ë
181	U+00CC	LATIN CAPITAL LETTER I WITH GRAVE	Ì
182	U+00CD	LATIN CAPITAL LETTER I WITH ACUTE	Í
183	U+00CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX	Î
184	U+00CF	LATIN CAPITAL LETTER I WITH DIAERESIS	Ï
185	U+00D0	LATIN CAPITAL LETTER ETH	Ð
186	U+00D1	LATIN CAPITAL LETTER N WITH TILDE	Ñ
187	U+00D2	LATIN CAPITAL LETTER O WITH GRAVE	Ò
188	U+00D3	LATIN CAPITAL LETTER O WITH ACUTE	Ó
189	U+00D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX	Ô
190	U+00D5	LATIN CAPITAL LETTER O WITH TILDE	Õ
191	U+00D6	LATIN CAPITAL LETTER O WITH DIAERESIS	Ö
192	U+00D8	LATIN CAPITAL LETTER O WITH STROKE	Ø
193	U+00D9	LATIN CAPITAL LETTER U WITH GRAVE	Ù
194	U+00DA	LATIN CAPITAL LETTER U WITH ACUTE	Ú
195	U+00DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX	Û
196	<U+00DB,U+0304>	LATIN CAPITAL LETTER U WITH CIRCUMFLEX WITH COMBINING MACRON	Û̄
197	U+00DC	LATIN CAPITAL LETTER U WITH DIAERESIS	Ü

Nr.	Codepoints	Name	Char
198	U+00DD	LATIN CAPITAL LETTER Y WITH ACUTE	Ý
199	U+00DE	LATIN CAPITAL LETTER THORN	Þ
200	U+00DF	LATIN SMALL LETTER SHARP S	ß
201	U+00E0	LATIN SMALL LETTER A WITH GRAVE	à
202	U+00E1	LATIN SMALL LETTER A WITH ACUTE	á
203	U+00E2	LATIN SMALL LETTER A WITH CIRCUMFLEX	â
204	U+00E3	LATIN SMALL LETTER A WITH TILDE	ã
205	U+00E4	LATIN SMALL LETTER A WITH DIAERESIS	ä
206	U+00E5	LATIN SMALL LETTER A WITH RING ABOVE	å
207	U+00E6	LATIN SMALL LETTER AE	æ
208	U+00E7	LATIN SMALL LETTER C WITH CEDILLA	ç
209	U+00E8	LATIN SMALL LETTER E WITH GRAVE	è
210	U+00E9	LATIN SMALL LETTER E WITH ACUTE	é
211	U+00EA	LATIN SMALL LETTER E WITH CIRCUMFLEX	ê
212	U+00EB	LATIN SMALL LETTER E WITH DIAERESIS	ë
213	U+00EC	LATIN SMALL LETTER I WITH GRAVE	ì
214	U+00ED	LATIN SMALL LETTER I WITH ACUTE	í
215	U+00EE	LATIN SMALL LETTER I WITH CIRCUMFLEX	î
216	U+00EF	LATIN SMALL LETTER I WITH DIAERESIS	ï
217	U+00F0	LATIN SMALL LETTER ETH	ð
218	U+00F1	LATIN SMALL LETTER N WITH TILDE	ñ
219	U+00F2	LATIN SMALL LETTER O WITH GRAVE	ò
220	U+00F3	LATIN SMALL LETTER O WITH ACUTE	ó
221	U+00F4	LATIN SMALL LETTER O WITH CIRCUMFLEX	ô
222	U+00F5	LATIN SMALL LETTER O WITH TILDE	õ
223	U+00F6	LATIN SMALL LETTER O WITH DIAERESIS	ö
224	U+00F8	LATIN SMALL LETTER O WITH STROKE	ø
225	U+00F9	LATIN SMALL LETTER U WITH GRAVE	ù
226	U+00FA	LATIN SMALL LETTER U WITH ACUTE	ú
227	U+00FB	LATIN SMALL LETTER U WITH CIRCUMFLEX	û
228	<U+00FB,U+0304>	LATIN SMALL LETTER U WITH CIRCUMFLEX WITH COMBINING MACRON	ũ
229	U+00FC	LATIN SMALL LETTER U WITH DIAERESIS	ü
230	U+00FD	LATIN SMALL LETTER Y WITH ACUTE	ý
231	U+00FE	LATIN SMALL LETTER THORN	þ
232	U+00FF	LATIN SMALL LETTER Y WITH DIAERESIS	ÿ
233	<U+00FF,U+0301>	LATIN SMALL LETTER Y WITH DIAERESIS WITH COMBINING ACUTE ACCENT	ÿ́
234	U+0100	LATIN CAPITAL LETTER A WITH MACRON	Ā
235	U+0101	LATIN SMALL LETTER A WITH MACRON	ā
236	U+0102	LATIN CAPITAL LETTER A WITH BREVE	Ă
237	U+0103	LATIN SMALL LETTER A WITH BREVE	ă
238	U+0104	LATIN CAPITAL LETTER A WITH OGONEK	Ą
239	U+0105	LATIN SMALL LETTER A WITH OGONEK	ą
240	U+0106	LATIN CAPITAL LETTER C WITH ACUTE	Ć
241	U+0107	LATIN SMALL LETTER C WITH ACUTE	ć
242	U+0108	LATIN CAPITAL LETTER C WITH CIRCUMFLEX	Ĉ
243	U+0109	LATIN SMALL LETTER C WITH CIRCUMFLEX	ĉ

Nr.	Codepoints	Name	Char
244	U+010A	LATIN CAPITAL LETTER C WITH DOT ABOVE	Ĉ
245	U+010B	LATIN SMALL LETTER C WITH DOT ABOVE	ĉ
246	U+010C	LATIN CAPITAL LETTER C WITH CARON	Č
247	<U+010C,U+0315>	LATIN CAPITAL LETTER C WITH CARON WITH COMBINING COMMA ABOVE RIGHT	Č̣
248	<U+010C,U+0323>	LATIN CAPITAL LETTER C WITH CARON WITH COMBINING DOT BELOW	Č̣̇
249	U+010D	LATIN SMALL LETTER C WITH CARON	č
250	<U+010D,U+0315>	LATIN SMALL LETTER C WITH CARON WITH COMBINING COMMA ABOVE RIGHT	č̣
251	<U+010D,U+0323>	LATIN SMALL LETTER C WITH CARON WITH COMBINING DOT BELOW	č̣̇
252	U+010E	LATIN CAPITAL LETTER D WITH CARON	Ď
253	U+010F	LATIN SMALL LETTER D WITH CARON	ď
254	U+0110	LATIN CAPITAL LETTER D WITH STROKE	Ɖ
255	U+0111	LATIN SMALL LETTER D WITH STROKE	ɖ
256	U+0112	LATIN CAPITAL LETTER E WITH MACRON	Ē
257	U+0113	LATIN SMALL LETTER E WITH MACRON	ē
258	U+0114	LATIN CAPITAL LETTER E WITH BREVE	Ĕ
259	U+0115	LATIN SMALL LETTER E WITH BREVE	ĕ
260	U+0116	LATIN CAPITAL LETTER E WITH DOT ABOVE	Ê
261	U+0117	LATIN SMALL LETTER E WITH DOT ABOVE	ê
262	U+0118	LATIN CAPITAL LETTER E WITH OGONEK	Ę
263	U+0119	LATIN SMALL LETTER E WITH OGONEK	ę
264	U+011A	LATIN CAPITAL LETTER E WITH CARON	Ě
265	U+011B	LATIN SMALL LETTER E WITH CARON	ě
266	U+011C	LATIN CAPITAL LETTER G WITH CIRCUMFLEX	Ĝ
267	U+011D	LATIN SMALL LETTER G WITH CIRCUMFLEX	ĝ
268	U+011E	LATIN CAPITAL LETTER G WITH BREVE	Ğ
269	U+011F	LATIN SMALL LETTER G WITH BREVE	ğ
270	U+0120	LATIN CAPITAL LETTER G WITH DOT ABOVE	Ġ
271	U+0121	LATIN SMALL LETTER G WITH DOT ABOVE	ġ
272	U+0122	LATIN CAPITAL LETTER G WITH CEDILLA	Ģ
273	U+0123	LATIN SMALL LETTER G WITH CEDILLA	ģ
274	U+0124	LATIN CAPITAL LETTER H WITH CIRCUMFLEX	Ĥ
275	U+0125	LATIN SMALL LETTER H WITH CIRCUMFLEX	ĥ
276	U+0126	LATIN CAPITAL LETTER H WITH STROKE	Ħ
277	U+0127	LATIN SMALL LETTER H WITH STROKE	ħ
278	U+0128	LATIN CAPITAL LETTER I WITH TILDE	Ĩ
279	U+0129	LATIN SMALL LETTER I WITH TILDE	ĩ
280	U+012A	LATIN CAPITAL LETTER I WITH MACRON	Ī
281	<U+012A,U+0301>	LATIN CAPITAL LETTER I WITH MACRON WITH COMBINING ACUTE ACCENT	Ī́
282	U+012B	LATIN SMALL LETTER I WITH MACRON	ī
283	<U+012B,U+0301>	LATIN SMALL LETTER I WITH MACRON WITH COMBINING ACUTE ACCENT	ī́
284	U+012C	LATIN CAPITAL LETTER I WITH BREVE	Ĭ
285	U+012D	LATIN SMALL LETTER I WITH BREVE	ĭ
286	U+012E	LATIN CAPITAL LETTER I WITH OGONEK	Į
287	U+012F	LATIN SMALL LETTER I WITH OGONEK	į

Nr.	Codepoints	Name	Char
288	U+0130	LATIN CAPITAL LETTER I WITH DOT ABOVE	İ
289	U+0131	LATIN SMALL LETTER DOTLESS I	ı
290	U+0132	LATIN CAPITAL LIGATURE IJ	IJ
291	U+0133	LATIN SMALL LIGATURE IJ	ij
292	U+0134	LATIN CAPITAL LETTER J WITH CIRCUMFLEX	Ĵ
293	U+0135	LATIN SMALL LETTER J WITH CIRCUMFLEX	ĵ
294	U+0136	LATIN CAPITAL LETTER K WITH CEDILLA	Ķ
295	U+0137	LATIN SMALL LETTER K WITH CEDILLA	ķ
296	U+0138	LATIN SMALL LETTER KRA	κ
297	U+0139	LATIN CAPITAL LETTER L WITH ACUTE	Ĺ
298	U+013A	LATIN SMALL LETTER L WITH ACUTE	ĺ
299	U+013B	LATIN CAPITAL LETTER L WITH CEDILLA	Ľ
300	U+013C	LATIN SMALL LETTER L WITH CEDILLA	ļ
301	U+013D	LATIN CAPITAL LETTER L WITH CARON	Ľ
302	U+013E	LATIN SMALL LETTER L WITH CARON	ľ
303	U+013F	LATIN CAPITAL LETTER L WITH MIDDLE DOT	Ľ
304	U+0140	LATIN SMALL LETTER L WITH MIDDLE DOT	ḷ
305	U+0141	LATIN CAPITAL LETTER L WITH STROKE	Ł
306	U+0142	LATIN SMALL LETTER L WITH STROKE	ł
307	U+0143	LATIN CAPITAL LETTER N WITH ACUTE	Ń
308	U+0144	LATIN SMALL LETTER N WITH ACUTE	ń
309	U+0145	LATIN CAPITAL LETTER N WITH CEDILLA	Ñ
310	U+0146	LATIN SMALL LETTER N WITH CEDILLA	ñ
311	U+0147	LATIN CAPITAL LETTER N WITH CARON	Ň
312	U+0148	LATIN SMALL LETTER N WITH CARON	ň
313	U+0149	LATIN SMALL LETTER N PRECEDED BY APOSTROPHE	’n
314	U+014A	LATIN CAPITAL LETTER ENG	Ŋ
315	U+014B	LATIN SMALL LETTER ENG	ŋ
316	U+014C	LATIN CAPITAL LETTER O WITH MACRON	Ō
317	U+014D	LATIN SMALL LETTER O WITH MACRON	ō
318	U+014E	LATIN CAPITAL LETTER O WITH BREVE	Ŏ
319	U+014F	LATIN SMALL LETTER O WITH BREVE	ö
320	U+0150	LATIN CAPITAL LETTER O WITH DOUBLE ACUTE	Ő
321	U+0151	LATIN SMALL LETTER O WITH DOUBLE ACUTE	ó
322	U+0152	LATIN CAPITAL LIGATURE OE	Œ
323	U+0153	LATIN SMALL LIGATURE OE	œ
324	U+0154	LATIN CAPITAL LETTER R WITH ACUTE	Ŕ
325	U+0155	LATIN SMALL LETTER R WITH ACUTE	ŗ
326	U+0156	LATIN CAPITAL LETTER R WITH CEDILLA	Ŗ
327	U+0157	LATIN SMALL LETTER R WITH CEDILLA	ŗ
328	U+0158	LATIN CAPITAL LETTER R WITH CARON	Ř
329	U+0159	LATIN SMALL LETTER R WITH CARON	ř
330	U+015A	LATIN CAPITAL LETTER S WITH ACUTE	Ś
331	U+015B	LATIN SMALL LETTER S WITH ACUTE	ś
332	U+015C	LATIN CAPITAL LETTER S WITH CIRCUMFLEX	Ŝ
333	U+015D	LATIN SMALL LETTER S WITH CIRCUMFLEX	ŝ
334	U+015E	LATIN CAPITAL LETTER S WITH CEDILLA	Ş

Nr.	Codepoints	Name	Char
335	U+015F	LATIN SMALL LETTER S WITH CEDILLA	ſ
336	U+0160	LATIN CAPITAL LETTER S WITH CARON	Š
337	U+0161	LATIN SMALL LETTER S WITH CARON	š
338	U+0162	LATIN CAPITAL LETTER T WITH CEDILLA	Ŧ
339	U+0163	LATIN SMALL LETTER T WITH CEDILLA	ŧ
340	U+0164	LATIN CAPITAL LETTER T WITH CARON	Ť
341	U+0165	LATIN SMALL LETTER T WITH CARON	ť
342	U+0166	LATIN CAPITAL LETTER T WITH STROKE	Ƨ
343	U+0167	LATIN SMALL LETTER T WITH STROKE	Ƨ
344	U+0168	LATIN CAPITAL LETTER U WITH TILDE	Ū
345	U+0169	LATIN SMALL LETTER U WITH TILDE	ū
346	U+016A	LATIN CAPITAL LETTER U WITH MACRON	Ū
347	U+016B	LATIN SMALL LETTER U WITH MACRON	ū
348	U+016C	LATIN CAPITAL LETTER U WITH BREVE	Ů
349	U+016D	LATIN SMALL LETTER U WITH BREVE	ů
350	U+016E	LATIN CAPITAL LETTER U WITH RING ABOVE	Ů
351	U+016F	LATIN SMALL LETTER U WITH RING ABOVE	ů
352	U+0170	LATIN CAPITAL LETTER U WITH DOUBLE ACUTE	Ű
353	U+0171	LATIN SMALL LETTER U WITH DOUBLE ACUTE	ű
354	U+0172	LATIN CAPITAL LETTER U WITH OGONEK	Ų
355	U+0173	LATIN SMALL LETTER U WITH OGONEK	ų
356	U+0174	LATIN CAPITAL LETTER W WITH CIRCUMFLEX	Ŵ
357	U+0175	LATIN SMALL LETTER W WITH CIRCUMFLEX	ŵ
358	U+0176	LATIN CAPITAL LETTER Y WITH CIRCUMFLEX	Ŷ
359	U+0177	LATIN SMALL LETTER Y WITH CIRCUMFLEX	ŷ
360	U+0178	LATIN CAPITAL LETTER Y WITH DIAERESIS	ÿ
361	U+0179	LATIN CAPITAL LETTER Z WITH ACUTE	Ž
362	U+017A	LATIN SMALL LETTER Z WITH ACUTE	ž
363	U+017B	LATIN CAPITAL LETTER Z WITH DOT ABOVE	Ẑ
364	U+017C	LATIN SMALL LETTER Z WITH DOT ABOVE	ẑ
365	U+017D	LATIN CAPITAL LETTER Z WITH CARON	Ž
366	<U+017D,U+0326>	LATIN CAPITAL LETTER Z WITH CARON WITH COMBINING COMMA BELOW	Ẓ
367	<U+017D,U+0327>	LATIN CAPITAL LETTER Z WITH CARON WITH COMBINING CEDILLA	Ẕ
368	U+017E	LATIN SMALL LETTER Z WITH CARON	ž
369	<U+017E,U+0326>	LATIN SMALL LETTER Z WITH CARON WITH COMBINING COMMA BELOW	ẓ
370	<U+017E,U+0327>	LATIN SMALL LETTER Z WITH CARON WITH COMBINING CEDILLA	ẕ
371	U+0187	LATIN CAPITAL LETTER C WITH HOOK	Ć
372	U+0188	LATIN SMALL LETTER C WITH HOOK	ć
373	U+018F	LATIN CAPITAL LETTER SCHWA	Ə
374	U+0197	LATIN CAPITAL LETTER I WITH STROKE	Ǝ
375	U+01A0	LATIN CAPITAL LETTER O WITH HORN	Ɔ
376	U+01A1	LATIN SMALL LETTER O WITH HORN	ɔ
377	U+01AF	LATIN CAPITAL LETTER U WITH HORN	Ʊ
378	U+01B0	LATIN SMALL LETTER U WITH HORN	ʊ
379	U+01B7	LATIN CAPITAL LETTER EZH	Ʒ
380	U+01CD	LATIN CAPITAL LETTER A WITH CARON	Ǻ

Nr.	Codepoints	Name	Char
381	U+01CE	LATIN SMALL LETTER A WITH CARON	ă
382	U+01CF	LATIN CAPITAL LETTER I WITH CARON	İ
383	U+01D0	LATIN SMALL LETTER I WITH CARON	ı
384	U+01D1	LATIN CAPITAL LETTER O WITH CARON	Ö
385	U+01D2	LATIN SMALL LETTER O WITH CARON	ö
386	U+01D3	LATIN CAPITAL LETTER U WITH CARON	Ů
387	U+01D4	LATIN SMALL LETTER U WITH CARON	ů
388	U+01D5	LATIN CAPITAL LETTER U WITH DIAERESIS AND MACRON	Ü
389	U+01D6	LATIN SMALL LETTER U WITH DIAERESIS AND MACRON	ü
390	U+01D7	LATIN CAPITAL LETTER U WITH DIAERESIS AND ACUTE	Ú
391	U+01D8	LATIN SMALL LETTER U WITH DIAERESIS AND ACUTE	ú
392	U+01D9	LATIN CAPITAL LETTER U WITH DIAERESIS AND CARON	Ů
393	U+01DA	LATIN SMALL LETTER U WITH DIAERESIS AND CARON	ů
394	U+01DB	LATIN CAPITAL LETTER U WITH DIAERESIS AND GRAVE	Ù
395	U+01DC	LATIN SMALL LETTER U WITH DIAERESIS AND GRAVE	ù
396	U+01DE	LATIN CAPITAL LETTER A WITH DIAERESIS AND MACRON	Ä
397	U+01DF	LATIN SMALL LETTER A WITH DIAERESIS AND MACRON	ä
398	U+01E2	LATIN CAPITAL LETTER AE WITH MACRON	Æ
399	U+01E3	LATIN SMALL LETTER AE WITH MACRON	æ
400	U+01E4	LATIN CAPITAL LETTER G WITH STROKE	Ɔ
401	U+01E5	LATIN SMALL LETTER G WITH STROKE	ɔ
402	U+01E6	LATIN CAPITAL LETTER G WITH CARON	Ǧ
403	U+01E7	LATIN SMALL LETTER G WITH CARON	ǧ
404	U+01E8	LATIN CAPITAL LETTER K WITH CARON	Ǭ
405	U+01E9	LATIN SMALL LETTER K WITH CARON	ǭ
406	U+01EA	LATIN CAPITAL LETTER O WITH OGONEK	Q
407	U+01EB	LATIN SMALL LETTER O WITH OGONEK	q
408	U+01EC	LATIN CAPITAL LETTER O WITH OGONEK AND MACRON	Ų
409	U+01ED	LATIN SMALL LETTER O WITH OGONEK AND MACRON	ų
410	U+01EE	LATIN CAPITAL LETTER EZH WITH CARON	Š
411	U+01EF	LATIN SMALL LETTER EZH WITH CARON	š
412	U+01F0	LATIN SMALL LETTER J WITH CARON	Ĳ
413	U+01F4	LATIN CAPITAL LETTER G WITH ACUTE	Ǵ
414	U+01F5	LATIN SMALL LETTER G WITH ACUTE	ǵ
415	U+01F8	LATIN CAPITAL LETTER N WITH GRAVE	Ñ
416	U+01F9	LATIN SMALL LETTER N WITH GRAVE	ñ
417	U+01FA	LATIN CAPITAL LETTER A WITH RING ABOVE AND ACUTE	Ą
418	U+01FB	LATIN SMALL LETTER A WITH RING ABOVE AND ACUTE	ą
419	U+01FC	LATIN CAPITAL LETTER AE WITH ACUTE	Ĳ
420	U+01FD	LATIN SMALL LETTER AE WITH ACUTE	ķ
421	U+01FE	LATIN CAPITAL LETTER O WITH STROKE AND ACUTE	Ø
422	U+01FF	LATIN SMALL LETTER O WITH STROKE AND ACUTE	ø
423	U+0212	LATIN CAPITAL LETTER R WITH INVERTED BREVE	Ŕ
424	U+0213	LATIN SMALL LETTER R WITH INVERTED BREVE	ŕ
425	U+0218	LATIN CAPITAL LETTER S WITH COMMA BELOW	Ș
426	U+0219	LATIN SMALL LETTER S WITH COMMA BELOW	ș
427	U+021A	LATIN CAPITAL LETTER T WITH COMMA BELOW	Ț